# WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



B15

### INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

**A1** 

(11) International Publication Number:

WO 99/51631

C07K 14/325

. .

(43) International Publication Date:

14 October 1999 (14.10.99)

(21) International Application Number:

PCT/F199/00290

(22) International Filing Date:

6 April 1999 (06.04,99)

(30) Priority Data:

980782

3 April 1998 (03.04.98)

FI

(71)(72) Applicants and Inventors: KORHONEN, Timo [FI/FI]; Huntutie 10 A, FIN-00950 Helsinki (FI). PALVA, Airi [FI/FI]; Välitalontie 87 A, FIN-00660 Helsinki (FI). PALVA, Ilkka [FI/FI]; Välitalontie 87 A, FIN-00660 Helsinki (FI). HYNÖNEN, Ulla [FI/FI]; Raisiontie 4 A 6, FIN-00280 Helsinki (FI). WESTERLUND-WIKSTRÖM, Benita [FI/FI]; Tähtelänkuja 5 A 3, FIN-10210 Inkoo (FI).

(74) Agent: SEPPO LAINE OY; Itämerenkatu 3 B, FIN-00180 Helsinki (FI).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent-(GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: A PROTEIN REGION RESPONSIBLE OF BINDING TO EPITHELIAL CELL TYPES AND A DNA SEQUENCE ENCODING SAID REGION

(57) Abstract

This invention relates to a DNA molecule encoding a polypeptide responsible of binding to human and/or animal epithelial cell types. It has been found that various fragments of S-layer protein SlpA of *Lactobacillus brevis* has adhesive properties to epithelial cells types. It is possible to modify or improve the binding capacity of various prokaryotic or eucaryotic cells to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cell types by using lactobacillar surface structures of this invention. In particular, it is possible with the nucleotide sequences of this invention to improve the binding properties of a host cell having probiotic effects to human and/or animal epithelial cell types.

# FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho		SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania		SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg		SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia		SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco		TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova		TG	Togo
BB	Barbados	GH	Ghana	MG.	Madagascar		TJ	Tajikistan
BE	Belgium	GN	Guinea	MK ·	The former Yugoslav		TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia		TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali		TT	Trinidad and Tobago
BJ	Benin	. IE	Ireland	MN	Mongolia		UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	.*.	UG	Uganda
BY	Belarus	IS .	Iceland	MW .	Malawi		US	United States of America
CA	Canada	IT	Italy	MX	Mexico .		UZ	Uzbekistan
CF	Central African Republic	JP `	Japan	NE	Niger		VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands		YU	Yugoslavia
CH	Switzerland	KG ·	Kyrgyzstan	NO	Norway		zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand			
СМ	Cameroon		Republic of Korea	PL	Poland			
CN	China	KR	Republic of Korea	PT ·	Portugal			
Cυ	Cuba	KZ	Kazakstan	RO	Romania			•
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation			
DE	Germany	LI	Liechtenstein	SD	Sudan			
DK.	Denmark	LK	Sri Lanka	SE	Sweden			
EE	Estonia	LR	Liberia	SG	Singapore		.:	• •

BNSDOCID: <WO\_\_\_9951631A1\_I\_>

CONTRACTOR CONTRACTOR

A PROTEIN REGION RESPONSIBLE OF BINDING TO EPITHELIAL CELL TYPES AND A DNA SEQUENCE ENCODING SAID REGION -

The present invention relates to DNA molecules encoding polypeptides responsible of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cells. Further this invention relates to vectors containing the DNA molecules and hosts transformed with the DNA molecules of this invention.

The present invention relates also to a method of constructing new hosts or new proteins

capable of binding to human and/or animal epithelial cell types.

This invention also relates to genes encoding preselected proteins modified to bind to human and/or animal epithelial cell types.

In addition this invention relates to a host cell, to a protein and to a method for carrying preselected factors/properties to human and/or animal epithelial cells or cell surfaces by using a *Lactobacillus brevis* strain or *L. brevis* SlpA protein.

Bacterial adhesion to human epithelial and subepithelial tissues is a decisive initial event in successful colonization of tissue sites by invading bacteria. Several molecular ligand-receptor interactions have been characterized for bacterial species that cause infectious diseases in man or animals. Adhesion of pathogenic bacteria to the tissue at the infection site helps the bacteria to resist mechanical defences of our body, such as peristalsis in the intestine or flow of urine in the urinary tract. Adhesion is a key determinant in the host, tissue and cell-type tropism of bacterial infections. Attachment to tissues is also important for those bacteria that establish themselves as members of the normal bacterial flora in the human body. Species of *Lactobacillus* are major members of the indigenous bacterial flora in the gastrointestinal and the genital tract of man and animals. Lactobacilli are thought to be beneficial to their host organism and have a long history of use in the gastrointestinal and

the urogenital tract to prevent or cure various minor illnesses. Probiotic effects of lactobacilli include exclusion of invasive pathogens from intestinal and vaginal surfaces, production of antimicrobial substances, stimulation of immune systems, as well as other physiological effects. As lactobacilli are members of the normal bacterial flora and foodgrade organisms, their possible use as carriers of vaccine antigens in the intestine has aroused interest. The mechanisms by which lactobacilli bring about the probiotic effects have remained uncharacterized, but it is generally agreed that efficient adhesion to epithelial surfaces is important for the colonization of the intestine as well as for the effects associated with these bacteria. Isolates of lactobacilli have been shown to adhere to the intestinal epithelium of their hosts (Coconnier et al., Appl. Environ. Microbiol 58:2034 - 2039 (1992) and Henriksson et al., Appl. Environ. Microbiol 57:499 - 502 (1991)), but to date, however, no molecular ligand-receptor interactions of lactobacilli have been characterized. Considering the high number of lactobacilli in our body and the biotechnological, health-associated as well as ecological importance of these bacteria, molecular characterization of their adhesion mechanisms is important.

S-layers are paracrystalline surface protein arrays that are commonly expressed by species of Eubacteria and Archaebacteria (reviewed in Messner and Sleytr, Adv. Microb. Physiol. 33: 213-275 (1992) and Sleytr and Sára, Trends Biotechnol. 15:20-26 (1997). Most S-layers are composed of a single protein species, the S-layer protein, greatly varying in size in different bacterial genera. The S-layer subunits are very hydrophobic and crystallize to form a two-dimensional layer on the bacterial surface. The genes encoding S-layers are efficiently transcribed, and the S-layer protein is the dominant protein species representing 10-20% of the total cellular protein of the bacterial cell. Differing functions have been attributed to the S-layers of different bacterial species. These functions include maintenance of the cell shape, protection of cells from hostile environment, anchorage of extracellular enzymes to the bacterial cell wall and mediation of bacterial attachment to animal tissues (Chu et al, J. Biol. Chem. 266: 15258-15265 (1991), Schneitz et al., J. Appl. Microbiol. 74: 290-294 (1993) and Toba et al., Appl. Environ. Microbiol. 61:2467-2471 (1995)). The S-layer of the fish pathogen Aeromonas salmonicida binds to extracellular matrix proteins and increases bacterial virulence by promoting bacterial spread to cause systemic infection in the

30

10

20

2.5

fish (Chu et al., J. Biol. Chem. 266: 15258-15265 (1991)). Most S-layer proteins aggregate in physiological buffers and their functional analysis have been restricted to solid phase assays (Sleytr and Sára, Trends Biotechnol. 15:20-26 (1997)), which has remained a severe limitation in the functional analysis of these important surface proteins.

- 5

200

15

20

S-layers are expressed by various species of the genus *Lactobacillus* (Masuda and Kawata *FEMS Microbiol. Lett.* 20:145-150 (1983). Their role in bacterial adhesiveness to chicken epithelium as well as to human and mouse extracellular matrix have been suggested (Schneitz *et al.*, *J. Appl. Microbiol.* 74: 290- 294 (1993), Toba *et al.*, *Appl. Environ. Microbiol.* 61:2467-2471 (1995)), but overall, the functions of lactobacillar S-layers have remained poorly characterized. Primary structure of a few lactobacillar S-layers have been determined (Vidgrén *et al. J. Bacteriol.* 174: 7419 - 7427 (1992), Boot *et al.*, *J. Bacteriol.* 175: 6089- 6096 (1993) and Boot *et al.*, *J. Bacteriol.* 177: 7222- 7230. The predicted lactobacillar S-layer proteins are 40 to 50 kDa in molecular size and show similarity in amino acid compositions.

Lactobacilli are important bacterial colonizers of our intestinal surfaces. Despite their high number and potential symbiotic effects in our body, our knowledge of the colonization mechanisms that the lactobacilli use to attach and multiply our intestine have remained uncharacterized. This has been in part due to the restricted methodology to genetically manipulate these bacteria and also due to the lack of suitable methods to study the binding mechanisms of these bacteria to intestinal or other mucosal surfaces.

WO 97/14802 suggests the use of Lactobacillus fermentum 104R 29 kD adherence factor for promoting the activity of microorganism cells to bind to a receptor recognized on mucus. However, the finding of a factor capable of binding to mucus does not solve the problem of specifically carrying preselected factors to human or animal epithelial cells.

Mucus on the surface of intestinal tract and any factors bound to the mucus are easily rinsed out from the intestine.

30

In this invention a DNA molecule encoding a protein region responsible of binding of the

protein to intestinal, urogenital, endothelial and/or other epithelial cell types has for the first time been identified and characterized. Although there has been some preliminary notions about the possible binding capability of the S-layer proteins of lactic acid bacteria (LAB) to human or animal cells, the binding property has not been confirmed to be due to the S-layer protein. The DNA molecule encoding a protein region responsible of binding was unexpectedly found from a gene encoding the S-layer protein of a species of lactic acid bacteria, *Lactobacillus brevis*. However, according to this invention a homologous DNA molecule encoding similar advantageous binding properties, may be synthetic or semisynthetic or originate from the same or another group of microorganisms.

10

15

This invention results in various advantages. This invention makes it for the first time possible to modify or improve the binding capacity of various prokaryotic or eukaryotic cells to human or animal epithelial cell types, like intestinal, urogenital and/or endothelial cell types by using lactobacillar surface structures. In particular, it is possible with the nucleotide sequences of this invention to improve the binding properties of a host cell having probiotic effects to human and/or animal epithelial cell types.

It is possible with the nucleotide sequences of this invention to modify or improve the binding properties of a preselected protein to human and/or animal epithelial cell types.

20

Furthermore it is possible by gene technological means to modify a host cell having or being modified to have the binding capability, to carry antigens, advantageously vaccine antigens to the intestinal and/or urogenital tract of humans and/or animals. In particular, it is possible to modify strains of *Lactobacillus brevis* species.

25

The probiotic properties of various strains of Lactobacillus brevis species can be improved by genetic modification. Totally new properties can be transferred to the strains of Lactobacillus brevis species or other hosts of this invention, which may carry these properties to human or animal gastrointestinal or urogenital tract.

30

Various host cells having or being modified to have the binding capability can be used to

colonize the human or animal gastrointestinal or urogenital tract and to exclude pathogens by binding to cell receptors, which would otherwise be bound by pathogens or by growing to cell surface areas which would otherwise be colonized by pathogens.

- Next, the invention will be described in more detail with the aid of the attached figures, of which
- Figure 1. Effect of S-layer removal on the adherence of L. brevis ATCC 8287 to human Intestine 407 cells. Panel a shows adherence of untreated bacteria to the intestinal cells, and panel b shows the adhesiveness of bacteria treated with guanidine hydrochloride to remove S-layer from the bacterial surface.
- Figure 2. Quantitative analysis of the effect of S-layer removal on the adherence of *L. brevis* to Intestine 407 cells. Panel A shows the number of adherent bacteria per epithelial cell; the adhesion test with the treated and the untreated bacteria was performed at four different bacterial cell densities indicated below. Panel B shows the polyacrylamide gel electrophoresis in sodium dodecyl sulfate (SDS-PAGE) of the proteins released from the bacterial surface by the Laemmli sample buffer used in SDS-PAGE. The S-layer peptide is indicated by the arrow. The S-layer peptide is the dominant peptide species released from the cells but not the only one.
- Figure 3. Schematic representation of the SlpA fragments expressed as fusion to flagellin. On top, the hydropathicity plot of the entire 465-amino acid SlpA peptide, The most probable antigenic determinant is indicated by the dashed line, and the SlpA peptide has a signal sequence of 30 amino acids. Below, the bars indicate the fragments expressed in *fliC*, the numbers refer to the N- and C-terminal amino acids in the SlpA peptide. Binding of the chimeric flagella to Intestine 407 (Int 407) cells is indicated on the right.
- Figure 4. Binding of the SlpA96-200/ FliC chimeric flagella to the human Intestine 407

  (Panel a) and urinary bladder T24 cells (Panel c). The binding was visualized by indirect immunofluorescence. Panels b and d show the corresponding fields by light microscopy.

Panel e shows the binding of the  $\Delta$  FliC flagella to Intestine 407 cells, and Panel g the control staining without flagella; the corresponding light microscopic fields are shown in Panels f and h. Arrows indicate binding of the chimeric flagella, arrowhead indicates lack of binding.

5

Figure 5. The nucleotide sequence of *Lactobacillus brevis SlpA* gene and the corresponding amino acid sequence. The first and last nucleotide residue of the coding sequence and of the various fragments is marked with an arrow. Correspondingly the first and last amino acid residue of the entire peptide and of the fragments is marked with a circle.

10

#### DNA molecules of the invention

In this invention a DNA molecule encoding a protein region responsible of binding to human and/or animal intestinal, urogenital, endothelial and/or other epithelial cell types has for the first time been identified and characterized.

In this invention it has been found that the S-layer protein SlpA of Lactobacillus brevis has adhesive properties to epithelial cell types. The primary structure of this protein and the corresponding gene has been described in WO 94/00581 and in Vidgren et al., J.

Bacteriol. 174: 7419-7427 (1992). According to Palva, A in H. Bahl et al./, FEMS
Microbiology Reviews 20 (1997):47 - 98 preliminary results have indicated that L. brevis S-layer protein could mediate binding to intestinal epithelial cells. However, until this date there has not been any suitable method to study the role of S-layer proteins of lactobacilli. For instance Schneitz et al. J. Appl. Microbiol. 74: 290- 294 (1993) have reported that the S-layer of L. acidophilus was involved in the adhesion of these bacteria to avian intestinal epithelial cells, but chemical removal of the S-layers did not affect the adhesion to Caco-2 cells. Guanidine hydrochloride extraction, which was used in the experiments, removes many protein types in addition to S-layer and hence the role of lactobacillar S-layers can not be deduced from these experiments.

30

. 25

By using a recently introduced flagella display system (Westerlund-Wikström et al. Protein

Engin. 10: 1319 - 1326 (1997)) we studied the adhesive properties of the S-layer protein SlpA of Lactobacillus brevis. We found, surprisingly, that the full length but also very short regions of the Lactobacillus brevis S-layer protein were capable of binding to epithelial cell types. The predicted size of the full-length S-layer protein is 465 amino acids. In the present invention various fragments of slpA were expressed as gene fusions in the variable region of the  $fliC_{H7}$  gene of Escherichia coli. The resulting chimeric flagella were assessed by indirect immunofluorescence for binding to various epithelial cells.

The S-layer peptides needed for the binding were contained in fragments of 270, 215, 275, 150 and 105 amino acid residues respectively, the shortest fragment representing residues 96 through 200 in the S-layer protein. However, any fragment being a partial amino acid sequence of these sequences or of the entire *Lactobacillus brevis* S-layer protein and possessing similar binding capacity as the above mentioned fragments is a polypeptide of this invention and any DNA molecule encoding these polypeptides is a DNA molecule of this invention.

Chimeric flagella harbouring inserts that represented the N-terminal part of the S-layer protein bound to intestinal as well as urinary bladder cells, whereas the C-terminal part of the S-layer protein did not confer binding on the chimeric flagella. However, the C-terminal parts of the protein may have an effect in enhancing the efficiency of binding.

The S-layer expressing bacterium *Lactobacillus brevis* ATCC 8287 efficiently adhered to the human small intestinal cell line and to the human urinary bladder cell line.

Bacterial adhesiveness to both cell lines was completely abolished after removal of the S-

layer protein (SlpA) from the bacterial surface by guanidine hydrochloride extraction.

This invention is directed to a DNA molecule encoding a polypeptide capable of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cells. The cells may originate from human or animal origin, like from porcine or poultry origin or from pet animals. The cells may be normal or e.g. tumour cells. The DNA molecule may be a DNA molecule having the full length or the partial sequence i.e. the coding sequence

30

contained in the nucleotide sequence of any one of SEQ ID NO. 1, SEQ ID NO. 2, SEQ ID NO. 3, SEQ ID NO. 4, SEQ ID NO 5 or SEQ ID NO. 6, representing the various fragments of the *Lactobacillus brevis slpA* gene, excluding, however, the full length SEQ ID NO. 6.

5

15

20

25

By a partial nucleotide sequence is meant a nucleotide sequence lacking at least one nucleotide residue as compared to the nucleotide sequences of SEQ ID 1 to SEQ ID 6.

SEQ ID NO. 1 represents the 315 nucleotide residues from 286 to 600, SEQ ID NO. 2 represents the 450 nucleotide residues from 286 to 735, SEQ ID NO. 3 represents the 825 nucleotide residues from 286 to 1110, SEQ ID NO. 4 represents the 645 nucleotide residues from 91 to 735, SEQ ID NO 5 represents the 810 nucleotide residues from 91 to 900 and SEQ ID NO. 6 represents the entire coding sequence of *Lactobacillus brevis slpA* gene from 1 to 1395 nucleotide residues. The first and last nucleotide residue of the nucleotide sequences of SEQ ID NO. 1 to SEQ ID NO. 6 is marked with an arrow in Figure 5.

The DNA molecule of this invention may be a DNA molecule encoding a polypeptide having the full length or the partial amino acid sequence i.e. the amino acid sequence contained in any one of SEQ ID NO. 7, SEQ ID NO. 8, SEQ ID NO. 9, SEQ ID NO.10, SEQ ID NO 11 or SEQ ID NO 12, representing the various fragments of *Lactobacillus brevis* SlpA protein, excluding, however, the full length SEQ ID NO. 12. SEQ ID NO. 7 represents the shortest amino acid sequence of 105 amino acid residues between 96 and 200, SEQ ID NO. 8 represents 150 amino acid residues between 96 and 245, SEQ ID NO. 9 represents 275 amino acid residues between 96 and 370, SEQ ID NO.10 represents 215 amino acid residues between 31 and 245 and SEQ ID NO 11 represents 270 amino acid residues between 31 and 300. SEQ ID NO 12 represents the amino acid sequence of the entire *Lactobacillus brevis* SlpA protein between 1 and 465. The first and last amino acid residue of the amino acid sequences of SEQ ID NO. 7 to SEQ ID NO. 12 is marked with a circle in Figure 5.

.30

10

By a partial amino acid sequence is meant an amino acid sequence lacking at least one amino acid compared to the amino acid sequences SEQ ID NO. 7 to SEQ ID NO. 12.

The present invention furthermore relates to DNA molecules, the sequences of which differ from the sequences of the above-identified molecules due to degeneracy of the genetic code, and which code for a polypeptide capable of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cells.

The present invention relates also to DNA molecules, the sequences of which hybridize to any one of the DNA molecules above encoding a polypeptide capable of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cells.

The term "hybridization" in this context means hybridization under conventional hybridization conditions, preferably under stringent conditions such as described by, e.g. Sambrook et al. (1989, Molecular Cloning, A Laboratory Manual 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Typical stringent hybridization conditions are exemplified in example 7, but equal hybridizations can be carried out in slightly different conditions as is known to a person skilled in the art.

- In example 7 the slpA gene of L. brevis has been hybridized to the chromosomal DNA of other Lactobacillus strains. As can be seen from example 7, the hybridization method is very useful and reliable method to find new DNA molecules of this invention. All the L. brevis strains tested gave positive hybridisation signal except two strains, which were shown to lack the S-layer protein. Other S-protein expressing lactobacilli belonging to other lactobacilli species gave negative result. L. buchneri, which is closely related to L. brevis gave also positive signal. L. buchneri carries S-layer and has the capability of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cells.
- 30 These nucleic acid molecules that hybridize to the nucleic acid molecules of the present

of Alfred St. Alto AND ADMINISTRATION

invention can in principle be derived from any organism possessing such nucleic acid molecules. Preferably, they are derived from lactic acid bacteria or bifidobacteria. Nucleic acid molecules hybridizing to the nucleic acid molecules of the present invention can be isolated, e.g., from genomic libraries or cDNA libraries of various organisms.

5

Such nucleic acid molecules can be identified and isolated by using the nucleic acid molecules of the present invention or fragments of these molecules or the reverse complements of these molecules, e.g. by hybridization according to standard techniques (see Sambrook *et al.*(1989)).

10

As hybridization probe, e.g. nucleic acid molecules can be used that have exactly or substantially the same nucleotide sequence indicated in the Figure 5 or fragments of said sequence. Preferably is used the entire nucleotide sequence of the coding sequence of the slpA gene. The fragments used as hybridization probes can also be synthetic fragments obtained by conventional synthesis techniques and the sequence of which is substantially identical to that of the nucleic acid molecules according to the invention. Once genes hybridizing to the nucleic acid molecules of the invention have been identified and isolated it is necessary to determine the sequence and to analyze the properties of the proteins coded for by said sequence.

20

The term "hybridizing DNA molecule" includes fragments, derivatives and allelic variants of the above-described nucleic acid molecules that code for the above-described protein (or its equivalent) or a biologically active fragment thereof. Fragments are understood to be parts of nucleic acid molecules long enough to code for the described protein (or its equivalent) or a biologically active fragment thereof. The term "derivative" means in this context that the nucleotide sequences of these molecules differ from the sequences of the above-described nucleic acid molecules in one or more positions and are highly homologous to said sequence.

The present invention is directed also to DNA molecules which are homologous with the DNA molecules contained in the coding sequences of any of SEQ ID 1 to SEQ ID 6 or with

the DNA molecules encoding a polypeptide having the amino acid sequence contained in any of SEQ ID 7 to SEQ ID 12 or with the degenerated forms of these DNA molecules "Homology" is understood to refer to a sequence identity of at least 50%, preferably more than 70% and still more preferably more than 90% on the length of at least 300 nucleotides. The deviations from the nucleic acid molecules described above can be the result of deletion, substitution, insertion, addition or combination.

Homology furthermore means that the respective nucleotide sequences or encoded proteins are functionally and/or structurally equivalent. The DNA molecules that are homologous to the DNA molecules described above and that are derivatives of said DNA molecules are regularly variations of said molecules which represent modifications having the same biological function. They may be naturally occurring variations, such as sequences of other organisms or mutations. These mutations may occur naturally or may be achieved by specific mutagenesis. Furthermore, these variations may be synthetically produced sequences.

The present invention furthermore relates to DNA molecules, the sequences of which have an amino acid sequence which shows at least 40 % identity, preferably at least 50 % identity, to a sequence contained above and which code for a polypeptide capable of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cells. The amino acid sequences of this invention show identity of less than 36% compared to any known amino acid sequence. When comparing the *L. brevis* slpA protein amino acid sequence with amino acid sequences of the prior art S-layer proteins known from lactobacilli, it was found that the highest identity was below 36 %.

25

15

20

By an amino acid sequence that is an "equivalent" of a specific amino acid sequence is meant an amino acid sequence that is not identical to the specific amino acid sequence, but rather contains at least some amino acid changes (deletion, substitutions, inversions, insertions, etc.) that do not essentially affect the biological activity of the protein as compared to a similar activity of the specific amino acid sequence, when used for a desired purpose. The biological activity of a polypeptide means here the capability of binding to epithelial cells.

Preferably, an "equivalent" amino acid sequence contains at least 40% - 99% identity at the amino acid level to the specific amino acid sequence, most preferably at least 50%, more preferably at least 60% and in an especially highly preferable embodiment, at least 95% identity, at the amino acid level.

5

The term "binding" is used here to mean the adherence of a cell, protein, protein region or polypeptide reasonably firmly to an epithelial cell, like intestinal, urogenital and/ or endothelial cell type. According to this invention the binding capacity has been measured by determining the binding of polypeptides encoded by DNA molecules expressed as gene fusions in the variable region of the  $fliC_{H7}$  gene of Escherichia coli. The resulting chimeric flagella were assessed by indirect immunofluorescence for binding to human intestinal and human urinary bladder cells. The binding capacity was visually characterized to be very strong (++++), strong (++++), weak (+) or no binding at all (-). By binding is meant the adherence of proteins or cells to the epithelial cells, particularly to the surface of the cell.

15

20

10

The binding was specific: chimeric flagella harbouring inserts that represented the N-terminal part of the S-layer protein bound to intestinal as well as urinary bladder cell types, whereas the C-terminal part of the S-layer protein did not confer binding on the chimeric flagella. The S-layer expressing bacterium *Lactobacillus brevis* ATCC 8287 efficiently adhered to the human small intestinal cell line and to the human urinary bladder cell line. Bacterial adhesiveness to both cell lines was completely abolished after removal of the S-layer protein (SlpA) from the bacterial surface by guanidine hydrochloride extraction.

In Table I is shown the differences in binding of different strains of lactic acid bacteria to human epithelial cell line (Intestine 407).

Table I. Binding of different strains of lactic acid bacteria to human epithelial cell line Intestine 407.

	Strains	•	Binding	Binding capacity							
5	L. acidophilus	JCM1132	<del>-</del>								
	L. crispatus J	CM5810	• .	+++*							
		A296-21	:	+							
	L. amylovorus	F81		· -							
0		JCM5807		+							
	L. gallinarum	T-50		-		· · .					
	L. gasseri	JCM1130		+ .							
	L. johnsonii	5F49		+							
	L. brevis	ATCC 8287		++++							
5		· · · · · · · · · · · · · · · · · · ·				,					

<sup>\*</sup> binding to the extracellular matrix secreted by the cell (see Toba et al., Appl. Environ. Microbiol. 61:2467-2471 (1995))

As can be seen in Table I, from the lactobacilli tested only L. brevis binds strongly to epithelial cells, particularly to the surface of the cell.

By "lactic acid bacteria" are meant all Gram-positive; anaerobic, microaerophilic or aerotolerant; catalase negative; rods or cocci; most importantly they all produce lactic acid as sole, major or important product from the energy-yielding fermentation of sugars. In practice, genuine members of lactic acid bacteria include at least the following genera: Lactobacillus, Lactococcus, Leuconostoc, Pediococcus, Carnobacterium, Sporolactobacillus, Streptococcus, Enterococcus, Aerococcus, Vagococcus, Tetragenococcus and Atopium. Many characteristics typical to genuine LAB are also common to the genus Bifidobacterium which consists of important health-promoting intestinal bacteria.

30

# Transfer of the binding property to new host cells

It is possible with the nucleotide sequences of this invention to modify or improve the binding capability of various prokaryotic or eukaryotic hosts to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cell types. The hosts cells are improved or modified to have the binding capability by transferring the host cells by at least one of the DNA molecules of this invention.

The binding capability may be transferred to any suitable bacterial host, for example to strains of lactic acid bacteria or bifidobacteria or to a fungal strain, like to a yeast strain.

A nucleotide sequence of this invention may be inserted into a DNA vector with conventional techniques, including blunt-ending or staggered-ending termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phophatase treatment to avoid undesirable joining, and ligation with appropriate ligases. Techniques for such manipulations are described in Sambrook et al. (1989, Molecular Cloning, A Laboratory Manual 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

To express a desired coding sequence, transcriptional and translational signals recognizable by the host are necessary. The nucleotide sequences of this invention may be operably linked to the transcriptional and secretory regulatory elements in an expression vector, and introduced into a host cell to produce desired protein under the control of such sequences. A DNA molecule is said to be capable of expressing a polypeptide if it contains expression control sequences which contain transcriptional regulatory information and such sequences are operably linked to the nucleotide sequence which encodes the desired polypeptide. An operable linkage is a linkage in which a sequence is connected to a regulatory sequence in such a way as to place expression of the sequence under the influence or control of the regulatory sequence.

30

15

Where the protein or protein region expression and secretion control sequences do not

function satisfactorily in a host cell, then sequences functional in the host cell may be substituted as necessary.

The vectors of the invention may comprise other operable linked functional elements such as DNA elements which confer antibiotic resistance on a host cell, and which provide for an origin of replication, or for insertion of a desired sequence into the chromosome of a host cell.

To transform a host cell with the DNA constructs of the invention many vector systems are available depending upon whether it is desired to insert the desired protein's DNA construct into the host cell chromosomal DNA, or to allow it to exist in an extrachromosomal form.

Cells which have stably integrated the introduced DNA into their chromosomes are selected by also introducing one or more markers which allow selection of host cells which contain the expression vector in the chromosome, for example the marker may provide resistance to antibiotics. The selectable marker gene (that can be later removed by methods well known in the art) can either be directly provided on the same vector as that providing the desired DNA gene sequences to be expressed, or such markers may be introduced into the same cell by co-transformation.

20

Factors of importance in selecting a particular plasmid or phage vector include the ease with which recipient cells that contain the vector may be recognized and selected from those recipient cells, which do not contain the vector and the number of copies of the vector which are desired in a particular host.

25

After the introduction of the vector, recipient cells are grown in a selective medium, which selects for the growth of vector-containing cells. Expression of the nucleotide sequences of this invention result in the production of the desired protein, or the production of a fragment of this protein and/or a host having the desired properties.

30

The above mentioned DNA molecules including the entire gene of L. brevis slpA can be

expressed on the surface of other selected well known probiotic microbes in order to enhance and target their adhesion to the host cells to allow more efficient colonization.

Also the degree of the adhesion mediated by the polypeptides of this invention can be varied by modifying their expression by different copy numbers, promoter strength regulation or by changing the length and amino acid composition of the binding domain with genetic engineering. With amino acid modifications encoded by the DNA molecules of this invention their host specificity may be affected if required.

10 Probiotics have been defined as live micro-organisms which beneficially affect the health of the host (human or animal) by improving its intestinal balance. To date, the term probiotic has been widened to include also live preparations used genitourinary to prevent infections and restore the disturbed microbiological ecological balance. Different characteristics associated with potentially health affecting bacteria may include e.g. i) acid and bile stability, ii) adherence to intestinal cells, iii) colonization of the intestinal tract, iv) production of antimicrobial substances, v) antagonism against pathogenic bacteria and vi) safety in food and clinical use.

According to this invention the probiotic effects of the hosts having (or modified to have)
the binding capacity to epithelial cells can be enhanced by genetic means.

Various hosts of this invention can be used to colonize the human or animal gastrointestinal or urogenital tract and to exclude pathogens by binding to cell receptors, which would otherwise be bound by pathogens or by growing to cell surface areas, which would otherwise be colonized by pathogens.

As an example of establishing a surface expression of the polypeptides of this invention, one can apply well conserved cell wall binding anchors (e.g. C-terminal sequence of PrtP from *L. lactis* or other lactic acid bacteria, staphylococcal protein A, *Streptococcus* pyogenes M6 protein, yeast anchor sequences etc.), fused to a spacer region and a DNA molecule of this invention which is preceded by a suitable promoter and signal sequence (see

30

e.g. Piard et al. J. Bacteriol. 179:3068-3072 (1997) and Steidler et al. Appl. Environ. Microbiol. 64: 342-345 (1998).

### Enhancing properties of oral vaccine carriers

5

The above mentioned DNA molecules including the entire gene of *L. brevis slpA* can be transferred to another microorganism which could be for example a live vaccine carrier or a new putative host developed for vaccine purpose to enhance their adhesion properties.

As described above the degree of adhesion can be varied by modifying the expression of the DNA molecules of this invention by different copy numbers, promoter strength regulation or by changing the length and amino acid composition of the binding domain with genetic engineering. With amino acid modifications encoded by the DNA molecules of this invention their host specificity may be affected if required.

15

20

25

The portals of entry of many pathogens are mucosal surfaces and most infections also begin at some mucosal site. The natural immune response to a pathogen is also likely to begin at the mucosal site of entry. To induce such a response with a vaccine, a similar route for its delivery may be regarded as a rational approach. Most vaccines available to date are, however, injected parenterally giving systemic response alone which may not be long lasting and protective enough. Oral vaccines are most likely to prevent intestinal diseases as they stimulate mucosal associated lymphoid tissue in the gastrointestinal tract directly. Thus, new strategies to prevent the adhesion, multiplication and invasion of pathogens at mucosal surfaces have become more acute and in some instances the only way to prevent an infectious disease. Furthermore, the use of oral (or other mucosal) routes for immunization against infective diseases is desirable because oral vaccines are easier to administer, have higher compliance rates and are likely to be less expensive. Experimental data indicate that colonization is required for example for immune response when using recombinant Streptococcus gordonii expressing surface antigens (Fischetti et al. Current Opinion in biotechnology 7:659-666 (1996) and Medaglini et al. Proc.Natl.Acad.Sci USA 92:6868-6872 (1995).

10

Mediating effective adherence to mucosal cells by the sequences of this invention, microorganisms carrying a homologous region to the identified binding region possess excellent properties as novel live oral vaccine carriers. Furthermore, it has been documented that surface (S-) layers function as effective adjuvants. In addition to the binding function, the C-terminal part of the SlpA gene can be modified to carry desired antigen epitopes and to efficiently present them as multiple copies at the S-layer formed by the SlpA subunits. Either a heterogenous or uniform S-layer can be formed depending on whether the host carries both an antigen-expressing and wild-type slpA gene or only the modified antigen-expressing gene. By this the amount of antigen on the cell surface may be affected. As described above, the vaccine antigen can also be expressed on the cell surface apart from a S-layer which in such a case functions as adhesion factor and adjuvant only.

Preferably expression hosts for oral vaccine carriers are strains of lactic acid bacteria and bifidobacteria, more preferably *L. brevis* species, in particular the strain *L. brevis*ATCC8287.

## Transfer of the binding property to new proteins

In addition to the use of the polypeptides of this invention responsible of the binding property in living vaccine carriers, also non-living applications for these structures are fully possible. The desired antigen polypeptides and shorter antigen epitopes having been modified to have the binding property of this invention can be expressed in the host chosen and administered after isolation and purification as pure protein preparations. If required, such preparations may be protected against degradation by using inert particles such as biodegradable microparticles, liposomes and cochelates (O'Hagan, *Novel Delivery Systems for Oral Vaccines*. CRC Press, Florida (1994)).

# Applications of the molecules of this invention in bifunctional target-specific biomolecules

An unexplored but highly possible application for the DNA molecules of this invention are their use in bifunctional molecules i.e. by creating fusion constructs where the binding region mediates the adhesion to mucosal cells and the other domain functions as an active target-specific molecule.

Such bifunctionally acting molecules could be formed for example by combining the DNA molecules of this invention with DNA sequences encoding enzymes, single chain antibodies or pharmaceutical proteins or toxins.

Targeted enzymes could be applied for example in the gut for degradation of (i) lactose by  $\beta$ -galactosidase to decrease adverse effects in lactose intolerance or ii) milk proteins by proteases and peptidases e.g. to release bioactive peptides or to increase tolerance against milk allergy. Certain toxins bound specifically to the polypeptides of this invention could also be used for destruction of pathogens at different mucosal sites. The delivery of the bifunctional molecule could be by direct spray or liquid preparations (nasal mucosa), by melting capsules (vaginal mucosa) or by inert particles described above for vaccine delivery systems (gastrointestinal tract).

With single chain antibodies and pharmaceutical proteins linked to the binding regions of this invention unlimited amounts of applications for mucosal prevention and medication of infectious diseases can be predicted.

# Use of Lactobacillus brevis as a probiotic host

According to this invention L. brevis as such may function as a novel probiotic strain particularly due to its highly efficient binding capacity both in the gut and urinary tract even though its use in practical applications is yet unraveled. Its probiotic effect can be further enhanced for example by introducing genes encoding i) production of selected

30

2.0

bacteriocin(s) and the corresponding immunity and ii) other antimicrobial substances to antagonist pathogens and iii) enzymes increasing its metabolic activity towards available substrates to strengthen its competetiveness in the chosen niche.

- Also the degree of the SlpA mediated adhesion can be varied by modifying its expression by different copy numbers, by promoter strength regulation or by changing the length and amino acid composition of the binding domain with genetic engineering. With amino acid modifications of the SlpA binding domain their host specificity may be affected if required.
- In addition to the demonstrated slpA encoded binding capacity of L. brevis ATCC8287, this strain is also highly resistant in low pH and bile thus possessing the key characteristics of a probiotic strain. L. brevis strains have also been shown to have antagonistic effects against intestinal pathogens. Furthermore, L. brevis can be regularly found in the intestine of man and animals as well as in a variety of fermented food products.

The following examples and figures provide further details of the invention.

## Example 1

15

Bacteria-The strain ATCC 8287 of L. brevis and its slpA gene have been described (Vidgrén et al. J. Bacteriol. 9:7419-7427 (1992)). The bacteria were grown overnight at 37 °C in 20 ml of static MRS broth. To extract the S-layer (Masuda, K. & Kawata, T. Microbiol. Immunol. 23:941-953 (1979)), the bacterial cells were washed once with phosphate buffered saline, pH 7.1 (PBS) and once with distilled water, and the cells were then suspended in 0.8 ml of distilled water. The cells were then incubated for 2 h at 37 °C with 7.2 ml of 2 M guanidine hydrochloride, the control cells were incubated similarly in PBS alone. The guanidine hydrochloride-treated cells were washed once with 2 M guanidine hydrochloride and then twice with PBS; the control cells were washed twice with PBS. For adherence assays, the bacterial cells were suspended in the cell culture medium used with the target epithelial cell lines (see below).

Bacterial adherence assays-The bacterial adhesion to cultured human epithelial cells was evaluated essentially as described earlier (Tarkkanen et al., Infect. Immunol. 65:1546-1549 (1997)). The human small intestine Intestine 407 (ATCC CCL6) cells were cultivated to confluence in RPMI 1640 medium (Gibco Laboratories, Grand Island, N.Y.) supplemented with 10 % (w/v) fetal calf serum (PAA Laboratories Gmbh, Linz, Austria), 1 % (w/v) Lglutamine (Life Technologies,), 1 % (w/v) nonessential amino acids (Gibco Laboratories, Grand Island, N.Y.), and gentamycin (50 µg/ml). The human urinary bladder transitional T24 cells (ATCC HTB-4) were cultered in McCoy's 5A medium (Life Technologies) supplemented with fetal calf serum, L-glutamine, and gentamycin as above. The cell lines were cultured on diagnostic glass slides (Knittel Glassbearbeitungs GmbH, Braunschweig, Germany). Before the adhesion assays, the cells were washed once with PBS. The bacteria were suspended in RPMI 1640 medium at the concentrations ranging from 5 x 10<sup>7</sup> to 10<sup>9</sup> cells/ml, and 40 µl of the suspension per well was added to the epithelial cells and the slides were incubated for 1 h at 37 °C in a moist chamber. The slides were washed five times at room temperature with PBS for five min each and fixed for 10 min with methanol. The cells with adherent bacteria were then examined in a BX50 microscope (Olympus Optical Co., Hamburg, Germany) either directly by Nomarski interference optics (for photographing) or, for quantitative analysis, stained for 5 min with 10 % v/v Giemsa stain and analyzed by light microscopy. The number and standard deviation of adherent bacteria on 20 epithelial cells was calculated. In inhibition studies, purified Fab fragments from anti-SlpA immunoglobulins (Vidgrén et al. J. Bacteriol. 9:7419-7427 (1992)) or from control immunoglobulins raised against fimbrial proteins of Escherichia coli (Westerlund et al. Zbl. Bakt. 278:229-237 (1993)) were incubated with the bacterial suspensions for 30 min over crushed ice prior to the adhesion assay, the control bacteria were incubated similarly but in PBS alone. The Fab fragments were prepared by a routine procedure (Porter, Biochem. J. 73:119-126 (1959)) and tested at the final concentration of 500  $\mu$ g/ml.

#### Example 2

Flagella display-The principle of the flagella display system used here was recently described (Westerlund-Wikström et al., Prot. Engin. 10:1319-1326 (1997)). Fragments

15

20

representing different parts of the slpA gene were amplified by polymerase chain reaction (PCR) with Pfu polymerase and using chromosomal DNA from the L. brevis strain ATCC 8287 as the template. The primers were designed on the basis of the nucleotide sequence of slpA (Vidgrén et al. J. Bacteriol. 9:7419-7427 (1992)) and contained an AccI restriction site at the 5' termini. The primers encoded the SIpA peptide sequences 31-245 (SEQ ID NO. 10, this peptide sequence is encoded by the nucleotide sequence SEQ ID NO. 4, which represents the nucleotide residues from 91-735), 31-300 (SEQ ID NO. 11 this peptide sequence is encoded by the nucleotide sequence SEQ ID NO. 5, which represents the nucleotide residues from 91-900), 96-200 (SEQ ID NO. 7; this peptide sequence is encoded by the nucleotide sequence SEQ ID NO. 1, which represents the nucleotide residues from 286-600), 96-245 (SEQ ID NO. 8; this peptide sequence is encoded by the nucleotide sequence SEQ ID NO. 2, which represents the nucleotide residues from 286-735), 96-370 (SEQ ID NO. 9; this peptide sequence is encoded by the nucleotide sequence SEQ ID NO. 3, which represents the nucleotide residues from 286-1110), or 239-447 (SEQ ID NO. 13; this peptide sequence is encoded by the nucleotide sequence SEQ ID NO. 14, which represents the nucleotide residues from 715-1341), where the residue numbers include the 30-mer signal sequence of the SlpA peptide. The slpA fragments were cloned into the AccI site in the plasmid pFliC<sub>H7\Delta</sub> deleted for 174 bp in the variable region of fliC and expressed in trans in E. coli that is fliC::Tn10 and fimA::cat (Westerlund-Wikström et al Prot. Eng. 10:1319-1326(1997)). The flagellar filaments were extracted and, after a sodium dodecylsulfate gel electrophoresis (SDS-PAGE), adjusted to an equal concentration of the FliC peptide as described recently (Westerlund-Wikström et al., Prot. Engin. 10:1319-1326 (1997)). The flagella lacking an insert, i.e. the ΔFliC filaments, were available from previous work (Westerlund-Wikström et al. (1997)).

25

15

#### Example 3

Binding tests with chimeric flagella-The binding of the chimeric flagella onto the epithelial cells was assessed by indirect immunofluorescence as detailed recently (Westerlund-Wikström et al. (1997)). Briefly, the epithelial cells were washed at room temperature with PBS, fixed with methanol for 10 min at -20 °C, and then washed with PBS at room

temperature. The flagellar extracts (40  $\mu$ l; 20  $\mu$ g/ml in PBS) were added and the slides were kept for 5 h at 4°C. After washing and a second fixing with methanol, the bound flagella were visualized by staining with immunoglobulin G molecules from an anti-H7-flagella rabbit antiserum and with fluorescein isothiocyanate-labelled secondary antibodies as detailed (Westerlund-Wikström *et al.*(1997)). The control assays included staining of the epithelial cells as above but using the  $\Delta$ FliC flagella lacking an insert, or omitting the flagellar extract, or the flagellar extract and the immunoglobulins in the staining procedure.

Immunological methods-For immuno electron microscopy, bacterial cells expressing the various flagellar constructs were suspended in Luria broth and immobilized on copper grids : 10 coated with Pioloform and carbon. The bacteria were left to react with an anti-H7-flagella antiserum (Westerlund-Wikström et al. (1997)), diluted 1/300 in PBS containing 10 mg/ml BSA or with an anti-SlpA antiserum (Vidgrén et al. J. Bacteriol. 9:7419-7427 (1992)); diluted 1/300 in PBS-BSA) for 90 min at room temperature. The grids were washed in PBS containing 1 mg/ml BSA, and the bound antibodies were detected with Auroprobe TMEM 15 Protein A-conjugate (Amersham, Amersham Place; Little Chalfont, Buckinghamshire, UK; diluted 1/40). The grids were examined in a Jeol JEM-100CX transmission electron microscope at an operating voltage of 60 kV. For immunoblotting, flagellar preparations were analyzed by SDS-PAGE using a 1 % (w/v) stacking gel and a 10 % separating gel. Polypeptides were transferred onto a nitrocellulose membrane using a semi-dry transfer apparatus at 0.9 mA/cm² membrane for 2 h at 4 °C. After transfer, the membrane was quenched with PBS containing 20 mg/ml BSA for 16 h at 20 °C and washed with PBS. Polypeptides were visualized by staining with diluted polyclonal anti-flagella antibodies or anti-SlpA antibodies and alkaline-fosfatase-conjugated secondary antibodies as described (Westerlund-Wikström et al. (1997)). A phosphatase substrate solution containing nitroblue-25 tetrazolium (162 μg/ml) and 5-bromo-4-chloro-3-indolyl-1-phosphate (370 μg/ml) was used.

#### Example 4

3.0

Adherence of L. brevis ATCC 8287 to intestinal cells-We initially assessed the adhesiveness

CAMPAGE CAN CLUSTED

of the *L. brevis* strain ATCC 8287 to the human small intestine cell line Intestine 407. The strain showed an efficient adhesion to the intestinal cells (Fig. 1A). The cells of *L. brevis* ATCC 8287 express the S-layer protein SlpA as their major cell surface protein (Vidgrén *et al. J. Bacteriol.* 9:7419-7427 (1992)), and we therefore extracted the S-layer from the bacterial surface and determined how this affected the adhesion. Extraction of cells with 2M guanidine hydrochloride is a routine procedure to remove bacterial S-layers, and the treatment does not lyse the bacterial cells. Treatment with guanidine hydrochloride completely abolished the adhesiveness of *L. brevis* (Fig. 1B). The S-layer peptide was the dominant peptide species released from the cells (Fig. 2B), and the results suggested a role for SlpA in bacterial adhesiveness.

#### Example 5

10

Expression of slpA fragments in fliC-The slpA gene encoding the S-layer protein of L. brevis ATCC 8287 has been described (Vidgrén et al. J. Bacteriol. 9:7419-7427 (1992)). 15 We cloned fragments of slpA into the AccI site in the plasmid pFliC<sub>H7A</sub> that contains a 174bp deletion in the variable region of fliC, and the chimeric flagella were expressed in E. coli JT1 that is fliC::Tn10 and fimA::cat. Schematic presentation of the slpA fragments expressed by the flagella display are shown in Fig.3. Western blots of the flagellar preparations with anti-H7 and anti-SlpA polyclonal antibodies showed that the apparent size 20 of the chimeric flagellins corresponded to those predicted from the nucleotide sequence, i.e. it increased with the size of the insert in  $\Delta fliC$  (data not shown). The polypeptides of smaller size that were present in the preparations and also reacted with the antibodies most likely were flagellar minor proteins (as noted earlier; Westerlund-Wikström et al., Prot. Engin. 10:1319-1326 (1997)) or degradation products. We noted by electron microscopy (data not 25 shown) that the chimeric flagella expressing the larger inserts (> 200 amino acids) had short flagellar filaments and thus these preparations were reduced in the relative amount of FliC as compared to the hook and cap proteins of the flagella. The chimeric flagella SlpA31-245/ΔFliC, SlpA31-300/ΔFliC, SlpA96-370/ΔFliC and SlpA239-447/ΔFliC reacted with both the anti-H7 and the anti-SlpA antibodies, whereas the flagella with the two shortest 30 inserts in the constructs SlpA96-245/ΔFliC and SlpA96-200/ΔFliC reacted with the anti-H7

antibodies and showed the expected apparent molecular size but failed to react with the anti-SlpA antiserum in Western blots.

#### Example 6

5..

Binding of chimeric flagella to epithelial cells-We next analyzed by an indirect immunofluorescence assay the binding of the chimeric flagella to Intestine 407 cells, representative examples of the assays are shown in Fig. 4. The SlpA31-300/ΔFliC construct, carrying the N-terminal part of the molecule, exhibited binding to the Intestine 407 cells, whereas the construct SlpA239-447/ΔFliC representing the C-terminal part, failed to bind (data not shown). No binding was observed with the ΔFliC construct lacking an insert. These results indicated that the binding region is located within the N-terminal part of the SlpA molecule. We constructed and tested various fragments covering different regions of the N-terminus, and the shortest fragment supporting adhesion to Intestine 407 cells was the construct SlpA96-200/ΔFliC that contained an 105 amino acid-long insert (Fig. 4A).

We also tested the binding of the SlpA96-200/ΔFliC construct to the human urinary bladder cell line T24 (Fig. 4C). The chimeric flagella bound efficiently to the urinary bladder cells, whereas no binding of the ΔFliC flagella were observed. The *L. brevis* ATCC 8287 cells exhibited an efficient adhesion to the urinary bladder cell line, and the adhesiveness was abolished after treatment of the bacteria with guanidine hydrochloride (data not shown). It was also ahown that *L. brevis* ATCC 8277 cells bound to CaCo-2 cells.

## Example 7

25

Hybridization of the slpA gene of L. brevis ATCC8287 to other Lactobacillus strainsThe aim of this experiment was to study the presence of L. brevis slp-gene homologs in other Lactobacillus strains. For this purpose chromosomal DNA was isolated from the following L. brevis strains: all the six strains available in the DSM culture collections (DSM 20054, 1267, 1268, 1269, 2647, 2647, 6235), five strains of the Japanese collection (Yasui et al. FEMS Microbiology Letter 133:183-186 (1995), one strain (VK3) from the culture

collection of TNO (delivered by prof. Peter Powels), one strain of the University of Cornell, U.S.A. (delivered by prof. Carl Batt) and one own isolate from pig intestine identified by API as L. brevis/L. buchneri. In addition, several chromosomal DNA samples of S-layer carrying Lactobacilli isolated from pig intestine and identified by API as putative members of L. acidophilus, L. fermentum, L. crispatus and L. delbrueckii species were tested.

Hybridization was performed in the following buffer: 5xSSC, 1% Blocking Reagent (Boehringer Mannheim), 0,02% SDS, 0,1% layryl sarcosine at 68°C overnight using the full length *L. brevis slpA* gene labelled with DIG (digoxigenin- dUTP, Boehringer Mannheim) as probe. Hybridization was followed by washing steps in descreasing salt concentration and increasing temperature. The final wash step was in 0.1xSSC, 0.1% SDS at 68°C for 15 min twice. The negative control used was and equal amount of *E.coli* or calf thymus DNA to that of test DNA.

All L. brevis strains tested gave a positive hybridization signal except the strain VK3 from TNO and the Japanese strain Yasui 0296961015. These two hybridization negative L. brevis strains were also shown to lack the S-layer protein by SDS-PAGE analysis. Furthermore, the other above mentioned S-layer expressing lactobacilli not belonging to the L. brevis and/or L. buchneri group remained negative in the hybridization test.

We have also shown previously in WO 94/00581 that the chromosomal DNA of *L. buchneri* strain DSM 20057 strongly hybridizes with the *slp*-gene probe used above under stringent conditions, whereas S-layer protein carrying *L. helveticus* and *L. acidophilus* strains remains hybridization negative under same conditions.

We have also shown that the *L. brevis/L. buchneri* strains hybridizing with the ATCC8287 slp-probe also effectively bind to the Intestine 407 cells whereas non-S-layer carrying *L. brevis* strains do not adhere to Intestine 407 cells.

BNSDOCID: «WO\_\_\_9951631A1\_I

20

## SEQUENCE LISTING

5	(1) GENERAL INFORMATION:	
10	<ul> <li>(i) APPLICANT:</li> <li>(A) NAME: Korhonen et al.</li> <li>(B) STREET: Huntutie 10 A</li> <li>(C) CITY: Helsinki</li> <li>(E) COUNTRY: Finland</li> <li>(F) POSTAL CODE (ZIP): 00950</li> </ul>	
15	(ii) TITLE OF INVENTION: A protein region responsible of binding epithelial cell types and a DNA sequence encoding said region	g to
	(iii) NUMBER OF SEQUENCES: 16	
20	<ul> <li>(iv) COMPUTER READABLE FORM:</li> <li>(A) MEDIUM TYPE: Floppy disk</li> <li>(B) COMPUTER: IBM PC compatible</li> <li>(C) OPERATING SYSTEM: PC-DOS/MS-DOS</li> <li>(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)</li> </ul>	
25	<pre>(vi) PRIOR APPLICATION DATA:   (A) APPLICATION NUMBER: FI 980782   (B) FILING DATE: 03-APR-1998</pre>	
**		
30	(2) INFORMATION FOR SEQ ID NO: 1:	
35	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 315 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: both  (D) TOPOLOGY: both	
	(ii) MOLECULE TYPE: DNA (genomic)	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
	GTTAAGACCA CTAACCGTGG TTCAGTTTAC TACCGTGTTG TAACGATGGA TGGCAAGTAC	60
·. 45	CGTGGTTACG TTTATGGTGG CAAGTCTGAC ACTGCCTTTG CTGGTGGTAT CAAGTCTGCT	120
	GAAACGACTA CTAAGGCTGA TATGCCTGCA CGTACTACTG GGTTCTACTT AACTGACACT	180
•	TCAAAGAACA CTCTTTGGAC GGCTCCTAAG TACACTCAAT ACAAGGCAAG TAAAGTTAGC	240
5 Q	CTTTATGGTG TTGCTAAGGA CACCAAGTTT ACTGTAGATC AGGCTGCTAC TAAGACTCGT	300
	GAAGGTTCAT TATAC	315
55	(2) INFORMATION FOR SEQ ID NO: 2:  (i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 450 base pairs	

(B) TYPE: nucleic acid(C) STRANDEDNESS: both

	(D) TOPOLOGY: both	
5	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
1.0	GTTAAGACCA CTAACCGTGG TTCAGTTTAC TACCGTGTTG TAACGATGGA TGGCAAGTAC	60
10	CGTGGTTACG TTTATGGTGG CAAGTCTGAC ACTGCCTTTG CTGGTGGTAT CAAGTCTGCT	120
	GAAACGACTA CTAAGGCTGA TATGCCTGCA CGTACTACTG GGTTCTACTT AACTGACACT	180
15	TCAAAGAACA CTCTTTGGAC GGCTCCTAAG TACACTCAAT ACAAGGCAAG TAAAGTTAGC	240
	CTTTATGGTG TTGCTAAGGA CACCAAGTTT ACTGTAGATC AGGCTGCTAC TAAGACTCGT	30
2.0	GAAGGTTCAT TATACTATCA CGTAACTGCT ACTAACGGTA GTGGTATTAG TGGTTGGATT	360
20	TACGCTGGTA AGGGCTTCAG TACTACTGCT ACTGGTACAC AAGTACTTGG TGGTCTGTCA	420
	ACTGATAAGT CAGTTACAGC AACCAACGAT	450
25 <sup>-</sup>	(2) INFORMATION FOR SEQ ID NO: 3:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 825 base pairs	•
•	(B) TYPE: nucleic acid	
30	(C) STRANDEDNESS: both (D) TOPOLOGY: both	-
	(ii) MOLECULE TYPE: DNA (genomic)	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:	
	GTTAAGACCA CTAACCGTGG TTCAGTTTAC TACCGTGTTG TAACGATGGA TGGCAAGTAC	61
	CGTGGTTACG TTTATGGTGG CAAGTCTGAC ACTGCCTTTG CTGGTGGTAT CAAGTCTGCT	120
40	GAAACGACTA CTAAGGCTGA TATGCCTGCA CGTACTACTG GGTTCTACTT AACTGACACT	180
	TCAAAGAACA CTCTTTGGAC GGCTCCTAAG TACACTCAAT ACAAGGCAAG TAAAGTTAGC	- 240
45	CTTTATGGTG TTGCTAAGGA CACCAAGTTT ACTGTAGATC AGGCTGCTAC TAAGACTCGT	300
	GAAGGTTCAT TATACTATCA CGTAACTGCT ACTAACGGTA GTGGTATTAG TGGTTGGATT	360
	TACGCTGGTA AGGGCTTCAG TACTACTGCT ACTGGTACAC AAGTACTTGG TGGTCTGTCA	420
50	ACTGATAAGT CAGTTACAGC AACCAACGAT AACAGTGTTA AGATTGTTTA CCGTACGACT	480
	GATGGCACTC AAGTTGGTTC TAACACTTGG GTAACTTCAA CTGATGGTAC AAAGGCAGGT	540
55	TCTAAGGTAA GCGATAAGGC CGCCGATCAA ACTGCTCTTG AAGCCTACAT CAATGCTAAC	600
	AAGCCTAGCG GTTACACTGT AACTAACCCT AATGCTGCAG ATGCTACCTA TGGTAACACA	660

	GTTTACGCTA CTGTTTCCCA AGCAGCTACT TCTAAGGTCG CTTTGAAGGT CTCAGGGACT	720
	CCTGTTACTA CTGCATTGAC TACAGCTGAT GCTAATGATA AGGTTGCAGC TAACGATACC	780
5	ACTGCTAATG GTAGTTCTGT TGCAGGCTCA ACAGTCTATG CTGCT	825
	(2) INFORMATION FOR SEQ ID NO: 4:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 645 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
	(C) STRANDEDNESS: both (D) TOPOLOGY: both	•
15	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:	
20	AAGTCATACG CTACTGCAGG TGCCTATTCA ACGTTAAAGA CGGACGCTGC TACTCGTAAC	60
	GTCGAAGCTA CTGGTACTAA CGCTTTATAC ACGAAGCCAG GTACTGTTAA GGGTGCTAAG	120
25	GTTGTCGCTT CTAAGGCTAC TATGGCTAAG TTAGCTTCTT CAAAGAAGTC AGCTGACTAC	180
20	TTCCGTGCTT ACGGTGTTAA GACCACTAAC CGTGGTTCAG TTTACTACCG TGTTGTAACG	240
	ATGGATGGCA AGTACCGTGG TTACGTTTAT GGTGGCAAGT CTGACACTGC CTTTGCTGGT	300
30	GGTATCAAGT CTGCTGAAAC GACTACTAAG GCTGATATGC CTGCACGTAC TACTGGGTTC	360
	TACTTAACTG ACACTTCAAA GAACACTCTT TGGACGGCTC CTAAGTACAC TCAATACAAG	420
35 :	GCAAGTAAAG TTAGCCTTTA TGGTGTTGCT AAGGACACCA AGTTTACTGT AGATCAGGCT	480
	GCTACTAAGA CTCGTGAAGG TTCATTATAC TATCACGTAA CTGCTACTAA CGGTAGTGGT	540
	ATTAGTGGTT GGATTTACGC TGGTAAGGGC TTCAGTACTA CTGCTACTGG TACACAAGTA	600
40	CTTGGTGGTC TGTCAACTGA TAAGTCAGTT ACAGCAACCA ACGAT	645
~	(2) INFORMATION FOR SEQ ID NO: 5:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 810 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: both</li></ul>	•
•	(D) TOPOLOGY: both	-
50	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:	
55∵	AAGTCATACG CTACTGCAGG TGCCTATTCA ACGTTAAAGA CGGACGCTGC TACTCGTAAC	60

GTCGAAGCTA CTGGTACTAA CGCTTTATAC ACGAAGCCAG GTACTGTTAA GGGTGCTAAG

	GTTGTCGCTT CTAAGGCTAC TATGGCTAAG TTAGCTTCTT CAAAGAAGIC AGCIGACIAC	100
	TTCCGTGCTT ACGGTGTTAA GACCACTAAC CGTGGTTCAG TTTACTACCG TGTTGTAACG	240
5	ATGGATGGCA AGTACCGTGG TTACGTTTAT GGTGGCAAGT CTGACACTGC CTTTGCTGGT	300
	GGTATCAAGT CTGCTGAAAC GACTACTAAG GCTGATATGC CTGCACGTAC TACTGGGTTC	360
10	TACTTAACTG ACACTTCAAA GAACACTCTT TGGACGGCTC CTAAGTACAC TCAATACAAG	420
	GCAAGTAAAG TTAGCCTTTA TGGTGTTGCT AAGGACACCA AGTTTACTGT AGATCAGGCT	480
	GCTACTAAGA CTCGTGAAGG TTCATTATAC TATCACGTAA CTGCTACTAA CGGTAGTGGT	540
15	ATTAGTGGTT GGATTTACGC TGGTAAGGGC TTCAGTÁCTA CTGCTACTGG TACACAAGTA	600
	CTTGGTGGTC TGTCAACTGA TAAGTCAGTT ACAGCAACCA ACGATAACAG TGTTAAGATT	.66L
20	GTTTACCGTA CGACTGATGG CACTCAAGTT GGTTCTAACA CTTGGGTAAC TTCAACTGAT	720
•	GGTACAAAGG CAGGTTCTAA GGTAAGCGAT AAGGCCGCCG ATCAAACTGC TCTTGAAGCC	780
	TACATCAATG CTAACAAGCC TAGCGGTTAC	810
25	(2) INFORMATION FOR SEQ ID NO: 6:	
30	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 1395 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: both</li> <li>(D) TOPOLOGY: both</li> </ul>	
	(ii) MOLECULE TYPE: DNA (genomic)	
35	(x) PUBLICATION INFORMATION: (A) AUTHORS: Vidgren, G	
<b>4</b> O	Palva, I Pakkanen, R Lounatmaa, K Palva, A	
15	(B) TITLE: S-Layer Protein Gene of Lactobacillus brevis:  Cloning by Polymerase Chain Reaction and  Determination of the Nucleotide Sequence	
ŧ⊃	(C) JOURNAL: J. Bacteriol. (D) VOLUME: 174 (E) ISSUE: 22 (F) PAGES: 7419-7427	
50	(G) DATE: 1992 (K) RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 1395	
55 ·	<ul> <li>(x) PUBLICATION INFORMATION:</li> <li>(H) DOCUMENT NUMBER: WO 94/00581 A1</li> <li>(I) FILING DATE: 24-JUN-1993</li> <li>(J) PUBLICATION DATE: 06-JAN-1994</li> </ul>	

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

	ATGCAATCAA	GTTTAAAGAA	ATCTCTTTAC	TTGGGCCTTG	CCGCATTGAG	CTTTGCTGGT	60
•	GTTGCTGCCG	TTTCAACGAC	TGCTTCAGCT	AAGTCATACG	CTACTGCAGG	TGCCTATTCA	120
5	ACGTTAAAGA	CGGACGCTGC	TACTCGTAAC	GTCGAAGCTA	CTGGTACTAA	CGCTTTATAC	180
	ACGAAGCCAG	GTACTGTTAA	GGGTGCTAAG	GTTGTCGCTT	CTAAGGCTAC	TATGGCTAAG	240
10	TTAGCTTCTT	CAAAGAAGTC	AGCTGACTAC	TTCCGTGCTT	ACGGTGTTAA	GACCACTAAC	300
	CGTGGTTCAG	TTTACTACCG	TGTTGTAACG	ATGGATGGCA	AGTACCGTGG	TTACGTTTAT	360
	GGTGGCAAGT	CTGACACTGC	CTTTGCTGGT	GGTATCAAGT	CTGCTGAAAC	GACTACTAAG	. 420
15	GCTGATATGC	CTGCACGTAC	TACTGGGTTC	TACTTAACTG	ACACTTCAAA	GAACACTCTT	480
	TGGACGGCTC	CTAAGTACAC	TCAATACAAG	GCAAGTAAAG	TTAGCCTTTA	TGGTGTTGCT	540
20	AAGGACACCA	AGTTTACTGT	AGATCAGGCT	GCTACTAAGA	CTCGTGAAGG	TTCATTATAC	600
	TATCACGTAA	CTGCTACTAA	CGGTAGTGGT	ATTAGTGGTT	GGATTTACGC	TGGTAAGGGC	660
	TTCAGTACTA	CTGCTACTGG	TACACAAGTA	CTTGGTGGTC	TGTCAACTGA	TAAGTCAGTT	720
25	ACAGCÁACCA	ACGATAACAG	TGTTAAGATT	GTTTACCGTA	CGACTGATGG	CACTCAAGTT	. 780
	GGTTCTAACA	CTTGGGTAAC	TTCAACTGAT	GGTACAAAGG	CAGGTTCTAA	GGTAAGCGAT	840
30	AAGGCCGCCG	ATCAAACTGC	TCTTGAAGCC	TACATCAATG	CTAACAAGCC	TAGCGGTTAC	900
	ACTGTAACTA	ACCCTAATGC	TGCAGATGCT	ACCTATGGTA	ACACAGTTTA	CGCTACTGTT	960
•	TCCCAAGCAG	CTACTTCTAA	GGTCGCTTTG	AAGGTCTCAG .	GGACTCCTGT	TACTACTGCA	1020
35	TTGACTACAG	CTGATGCTAA	TGATAAGGTT	GCAGCTAACG	ATACCACTGC	TAATGGTAGT	1080
	TCTGTTGCAG	GCTCAACAGT	CTATGCTGCT	GGTACTAAGT	TGGCTCAATT	AACAACTGAC	1140
40	TTGACTGGTG	AAAAGGGTCA	AGTTGTCACA	TTAACTGCCA	TCGATACTGA	TTTGGAAGAC	1200
	GCTACGTTCA	CTGGAACTAC	GACTTACTAT	TCAGATCTTG	GTAAAGCATA	CCACTACACT	1260
	TACACTTACA	ATAAGGACAG	TGCTGCTTCT	TCAAATGCAA	GTACCCAATT	TGGTTCAAAC	1320
45	GTCACTGGTA	CTTTAACTGC	FACCCTTGTT	ATGGGTAAGT	CTACTGCTAC	TGCTAACGGT	1380
	ACTACTTGGT	TCAAC					1395

(2) INFORMATION FOR SEQ ID NO: 7:

50

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 105 aminò acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

	(xi)	SEQU	JENCE	DES	CRI	OIT	 1: SE	EQ II	NO:	7:		₩.					
_		Val 1	Lys	Thr	Thr	Asn 5	Arg	Gly	Ser	Val	Tyr 10	Tyr	Arg	Val	Val	Thr 15	Met
		Asp	Gly	Lys	Tyr 20	Arg	Gly	Tyr	Val	Tyr 25	Gly	Gly	Lys	Ser	Asp 30	Thr	Ala
10		Phe	Ala	Gly 35	Gly	Ile	Lys	Ser	Ala 40	Glu	Thr	Thr	Thr	Lys 45	Ala	Asp	Met
		Pro	Ala 50	Arg	Thr	Thr	Gly	Phe 55	Tyr	Leu	Thr	Asp	Thr 60	Ser	Lys	Asn	Thr
15		Leu 65	Trp	Thr	Ala	Pro	Lys	Tyr	Thr	Gln	Tyr	Lýs 75	Ala	Ser	Lys	Val	Ser 80
2.0		Leu	Tyr	Gly	Val	Ala 85	Lys	Asp	Thṛ	Lys	Phe 90	Thr	Val	Asp	Gln	Ala 95	Ala
20		Thr	Lys	Thr	Arg 100	Glu	Gly	Ser	Leu	Tyr 105	. ,	ı				-	
	(2)	INFO	RMAT	ION I	FOR S	SEQ :	ID NO	D: 8	•	•							
25		(i)	SEQ	JÉNCI	E CH	ARAC'	reri:	STICS	· . 5 :								. • •
	× .					: 150 amino			cids	Š				•	· .		
30					RANDI POLO	EDNE:											
30	•	(ii)	(D)	) . TOI	50TÓ	<b>GY:</b> 1	both	ide		٠.					*		
30		(ii) (xi)	(D)	) TOI	POLO	GY: 1 PE: 1	both pept:		EQ II	o no	: 8:						
30 35		(xi)	(D) MOLI	) TOI	POLOG E TYI E DE:	GY: ] PE: ] SCRI	both pept: PTION	<b>1:</b> S		•		Tyr	Arg	Val	Val	Thr	Met
		(xi)	(D) MOLI	) TOI	POLOG E TYI E DE:	GY: ] PE: ] SCRI	both pept: PTION	<b>1:</b> S		O NO Val		Tyr	Arg	Val	Val	Thr 15	Met
		(xi) Val	(D) MOLI	) TOI ECULI UENCI Thr	POLOGE TY	PE: ] SCRI Asn	pept: PTION .Arg	N: SI	Ser	Val	Tyr 10					15	Met Ala
35		(xi) Val 1 Asp	(D) MOLI SEQU	ECULI ECULI UENCI Thr	POLOGE TYPE Thr Tyr 20	GY: 1 PE: 1 SCRI Asn 5 Arg	pept: PTION Arg	N: SI Gly Tyr	Ser Val	Val Tyr 25	Tyr 10 Gly	Gly	Lys	Ser	Asp 30	15 Thṛ	
35		(xi) Val 1 Asp	MOLI SEQU Lys Gly	ECULE UENCE Thr Lys Gly 35	Thr Tyr 20 Gly	GY: DE: DE: DE: DE: DE: DE: DE: DE: DE: DE	pept: PTION Arg Gly Lys	Gly Tyr Ser	Ser Val Ala	Val Tyr 25 Glu	Tyr 10 Gly Thr	Gly	Lys	Ser Lys 45	Asp 30 Ala	Thr Asp	Ala
35 40 45		(xi) Val 1 Asp Phe	MOLI SEQU Lys Gly Ala Ala 50	ECULIUENCE Thr Lys Gly 35 Arg	Thr  Gly  Thr	GY: PE: PE: PE: PE: PE: PE: PE: PE: PE: PE	pept: PTION Arg Gly Lys Gly	Gly Tyr Ser Phe	Ser Val Ala 40 Tyr	Val Tyr 25 Glu Leu	Tyr 10 Gly Thr	Gly Thr Asp	Lys Thr Thr	Ser Lys 45 Ser	Asp 30 Ala Lys	Thr Asp Asn	Ala Met
35		(xi) Val 1 Asp Phe Pro Leu 65	MOLI SEQU Lys Gly Ala Ala 50	ECULIUENCE Thr Lys Gly 35 Arg	Thr Tyr 20 Gly Thr	GY: PE: PE: PE: PE: PE: PE: PE: PE: PE: PE	pept: PTION Arg Gly Lys Gly Lys 70	Gly Tyr Ser Phe 55	Ser Val Ala 40 Tyr	Val Tyr 25 Glu Leu	Tyr 10 Gly Thr Thr	Gly Thr Asp Lys	Lys Thr Thr 60 Ala	.Ser Lys 45 Ser Ser	Asp 30 Ala Lys	Thr Asp Asn Val	Ala Met Thr Ser
35 40 45		(xi) Val 1 Asp Phe Pro Leu 65 Leu	MOLI SEQU Lys Gly Ala Ala 50 Trp	ECULE UENCE Thr Lys Gly 35 Arg Thr	Thr Tyr 20 Gly Thr Ala	GY: PE: PE: PE: PE: PE: PE: PE: PE: PE: PE	pept: PTION Arg Gly Lys Gly Lys 70 Lys	Gly Tyr Ser Phe 55 Tyr Asp	Ser Val Ala 40 Tyr Thr	Val Tyr 25 Glu Leu Gln Lys	Tyr 10 Gly Thr Thr Tyr	Gly Thr Asp Lys 75 Thr	Lys Thr Thr 60 Ala Val	Lys 45 Ser Ser	Asp 30 Ala Lys Lys	Thr Asp Asn Val Ala 95	Ala Met Thr Ser

									-	, ,							
				115					120					125			
5		Thr	Ala 130	Thr	Gly	Thr	Gln	Val 135	Leu	Gly	Gly		Ser 140	Thr	Asp	Lys	Ser
5		Val 145	Thr	Ala	Thr	Asn	Asp 150						_				
10	.(2)	INFO	RMAT	ION I	FOR S	SEQ :	ID NO	D: 9	:				-				
		(i)	(A (B (C	UENCI ) LEI ) TYI ) STI	NGTH PE: a RANDI	: 27! amino EDNE:	5 am: o ac: SS:	ino a		5							
15				) TO													
			•	ECULI			_				0						•
20		(xı)	SEQ	UENCI	: DES	SCRI	b.T.TOI	N: 51	≝Q II	טאָ כ	: 9:				. ,		
20		Val	Lys	Thr	Thr	Asn 5	Arg	Gly	Ser	Val	Tyr 10	Tyr	Arg	Val	Val	Thr 15	Met
25	٠.	Asp	Gly	Lys	Tyr 20	Arg	Gly	Tyr	Val	Tyr 25	Gly	Gly.	Lys	Ser	Asp 30	Thr	Ala
. •		Phe	Ala	Gly 35	Gly	Ile	Lys	Ser	Ala 40	Glu	Thr	Thr	Thr	Lys 45	Ala	Asp	Met
3.0		Pro	Ala 50	Arg	Thr	Thr	Gly	Phe 55	Tyr	Leu	Thr	Asp	Thr 60	Ser	Lys	Asn	Thr
35		Leu 65	Trp	Thr	Ala	Pro	Lys 70	Tyr	Thr	Gln	Tyr	Lys 75	Ala	Ser	Lys	Val	Ser 80
		Leu	Tyr	Gly	Val	Ala 85	Lys	Asp	Thr	Lys	Phe 90	Thr	Val	Asp	Gln	Ala 95	Ala
40		Thr	Lys	Thr	Arg 100	Glu	Gly	Ser	Leu	Tyr 105	Tyr	His	Val	Thr	Ala 110	Thr	Asn
		Gly	Ser	Gly 115	Ile	Ser	Gly	Trp	11e 120	Tyr	Ala	Glý	Lys	Gly 125	Phe	Ser	Thr
45		Thr	Ala 130	Thr	Gly	Thr	Gln	Val 135	Leu	Gly	Gly	Leu	Ser 140	Thr	Asp	Lys	Ser
50		Val 145	Thr	Ala	Thr	Asn	Asp 150	Asn	Ser	Val	Lys	Ile 155	Val	Tyr	Arg	Thr	Thr 160
		Asp	Gly	Thr	Gln	Val 165	Gly	Ser	Asn	Thr	Trp 170	Val	Thr	Ser	Thr	Asp 175	Gly
55	,	Thr	Lys	Ala	Gly 180	Ser	Lys	Val	Ser	Asp 185	Lys	Ala	Ala	Asp	Gln 190	Thr	Ala
		Leu	Glu	Ala	Tyr	Ile	Asn	Ala	Asn	Lys	Pro	Ser	Gly	Tyr	Thr	Val	Thr

			•						-	94								
				195					200					205				
5		Asn	Pro 210	Asn	Ala	Ala	Asp	Ala 215	Thr	Tyr	Gly	Asn	Thr 220	Val	Tyr	Ala	Thr	
		Val 225	Ser	Gln	Ala	Ala	Thr 230	Ser	Lys	Val	Ala	Leu 235	Lys -	Val	Ser		Thr. 240	
10		Pro	Val	Thr	Thr	Ala 245	Leu	Thr	Thr	Ala	Asp 250	Ala	Asn	Asp	Lys	Val 255	Ala	
	٠.	Ala	Asn	Asp	Thr 260	Thr	Ala	Asn	Gly	Ser 265	Ser	Val	Ala	Gly	Ser 270	Thr	Val	
15		Tyr	Ala	Ala 275	• .				•									
	(2)	INFO	RMAT:	ION 1	FOR S	SEQ :	ID NO	D: 10	):									
20 .		(i)	(A) (B) (C)	UENCI ) LEI ) TYI	NGTH PE: 8 RANDI	: 219 amino EDNE	5 am: o ac: SS:	ino, a										
25		*	(D)	) TO	FOTO	3Y: 1	ocn								٠			٠
		(ii)	MOL	ECULI	E TYI	PE: 1	pept	ide				. *						
•		(xi)	SEQUI	ENCE	DES	CRIP	rion	: SE(	O ID	NO:	10:	. •						
30		Lys.	Ser	Tyr	Ala	Thr 5	Ala	Gly	Ala	Tyr	Ser 10	Thr	Leu	Lys	Thr	Asp 15	Ala	
35		Ala	Thr	Arg	Asn 20	Val	Glu	Ala	Thr	Gly 25	Thr	Asn	Ala	Leu	Tyr 30	Thr	Lys	
		Pro	Gly	Thr 35	Val	Lys	Gly	Ala	Lys 40	Val	Val	Ala	Ser	Lys 45	Ala	Thr	Met	
40			50					55					60				Tyr	•
		Gly 65	Val	Lys	Thr	Thr	Asn 70	Arg	Gly	Ser	Val	Tyr 75	Tyr	Arg	Val	Val.	Thr 80	
45		Met	Asp	Gly	Lys	Tyr 85	Arg	Gly	Tyr	Val	Tyr 90	Gly	Gly	Lys		Asp 95	Thr	
50		Ala	Phe	Ala	Gly 100	Gly	Ile	Lys	Ser	Ala 105	Glu	Thr			Lys 110	Ala	Asp	
		Met	Pro	Ala 115	Arg	Thr	Thr	Glý	Phe 120		Leu	Thr	Asp	Thr 125	Ser	Lys	Asn	
55	. •	Thr	Leu 130		Thr	Ala	Pro	Lys 135	_	Thr	Gln	Tyr	Lys 140	Ala	Ser	Lys	Val	
		Ser	Leu	Tyr	Gly	Val	Ala	Lys	Asp	Thr	Lys	Phe	Thr	Val	Asp	Gln	Ala	

													•				
		145	5				150	)				15	5		٠.		160
, 5		Ala	a Thi	Lys	s Thi	169		ı Gly	/ Sei	r Let	1 Tyr 170		r His	s Va.	l Thi	Ala 175	
		Ası	ı Gly	/ Sei	r Gly 180		e Sei	c Gly	Trp	11e	_	Ala	a Gly	y Lys	5 Gl <sub>}</sub>		e Ser
10		Thi	Thi	Ala 199		Gly	/ Thi	Glr.	Va]		ı Gly	gl <sub>y</sub>	/ Let	1 Se1 205		: Asp	Lys
		Ser	Val 210		Ala	Thr	Asr	Asp 215									
15	(2)	INFO	RMAT	CION	FOR	SEQ	ID N	<b>I</b> O: 1	1:								
٠		(i)	( <i>P</i>	) LE		: 27	'0 aπ	STIC		ls	·			•			
20			(C	) SI	RAND	EDNE	:SS:			•							
	•	(ii)	MOL	ECUL	E TY	PE:	pept	ide								٠	٠.
25		(xi)	SEQ	UENC	E DE	SCRI	PTIO	N: S	EQ I	р ио	: 11	:					
30		Lys 1	Ser	Týr	Ala	Thr 5	Ala	Gly	Ala	Tyr	Ser 10	Thr	Leu	Lys	Thr	Asp 15	Ala
	٠.	Ala	Thr	Arg	Asn 20	Val	Glu	Ala	Thr	Gly 25	Thr	Asn	Ala	Leu	Tyr 30	Thr	Lys
35		Pro	Gly	Thr	Val	Lys	Gly	Ala	Lys 40	Val	Val	Ala	Ser	Lys 45	Ala	Thr	Met
		Ala		Leu	Ala	Ser	Ser	Lys 55	Lys	Ser	Ala	Asp	Tyr 60	Phe	Arg	Ala	Tyr
40		Gly 65	Val	Lys	Thr	Thr	Asn 70	Arg	Gly	Ser	Val	Tyr 75	Tyr	Arg	Val	Val	Thr 80
45		Met	Asp	Gly	Lys	Tyr 85	Arg	Gly	Tyr	Val	Tyr 90	Gly	Gly	Lys	Ser	Asp 95	Thr
		Ala	Phe	Ala	Gly 100	Gly	Ile	Lys	Ser	Ala 105	Glu	Thr	Thr	Thr	Lys 110	Ala	Asp
50		Met	Pro	Ala 115	Arg	Thr	Thr	Gly	Phe 120	Tyr	Leu	Thr	Asp	Thr 125	Ser	Lys	Asn
	•	Thr	Leu 130	Trp	Thr	Ala	Pro	Lys 135	Tyr.	Thr	Gln	Tyr	Lys 140	Ala	Ser	Lys	Val
55		Ser	Leu	Tyr	Gly	Val	Ala	Г'nз	Asp	Thr	Lys	Phe	Thr	Val	Asp	Gln	Ala

									٠.								
		Ala	Thr	Lys	Thr	Arg 165	Glu	Gly	Ser	Leu	Tyr 170	Tyr	His	Val	Thr	Ala 175	Thr
5		Asn	Gly	Ser	Gly 180	Ile	Ser	Gly	Trp	Ile 185	Tyr	Ala	Gly	Lys	Gly 190	Phe	Ser
		Thr	Thr	Ala 195	Thr	Gly	Thr	Gln	Val 200	Leu	Gly	Gly	Leu	-Ser 205	Thr	Asp	Lys
10		Ser	Val- 210	Thr	Ala	Thr	Asn	Asp 215	Asn	Ser	Val	Lys	Ile 220	Val	Tyr	Arg	Thr
		Thr 225	Asp	Gly	Thr	Gln	Val 230	Gly	Ser	Asn	Thr	Trp 235	Val	Thr	Ser	Thr	Asp 240
15	٠.	Gly	Thr	Lys	Ala	Gly 245	Ser	Lys	Val	Ser	Asp 250	Lys	Ala	Ala	Asp	Gln 255	Thr
20	. (1)	Ala	Leu	Glu	Ala 260	Tyr	Ile	Asn	Ala	Asn 265	Lys	Pro	Ser	Gly	Tyr 270		
	(2)	INFO	RMAT	ION I	FOR S	SEQ :	ID NO	D: 12	2:								
							·		· ·			•		-			
25		(1)	(A)	) LE	E CHA NGTH PE: a	: 469	am:	ino a	s: acids	3							
					RANDI POLO					• •	•			• .			
30		(ii)	WOLI	ECUL	E TY	PE: p	prot	ein	•								
		(x)			rion rhors	s: V:	idgr	en, (		-							
<b>ر</b> د							alva										
35								nen, cmaa									
				. 1			alva		, 1		•						•
			(B)												ıs bı	revis	<b>:</b>
40														n and			
40			(C)	זסד, ו	JRNAJ					з ипс	steo	TIGE	seq	Telice	5	•	
					LUME				:								
					SUE:									•			
45					GES:		9-742	27									
40					re: : Levai		ESIDU	JES :	IN SI	EO II	ONO:	12:	: FRO	OM 1	TO 4	165	
	٠.		•							-							
		(x)			rion							_		•			
50	•								) 94, N-199		31 A	L					
									06-J <i>I</i>		994						
		(xi)	SEQ	JENCI	E DES	SCRII	PTIO	1: SI	EQ II	ONO:	: 12	•					
55		Met 1	Gln	Ser	Ser	Leu 5	Lys	Lys	Ser	Leu	Tyr 10	Leu	Gly	Leu	Ala	Ala 15	Leu <sub>.</sub>

					٠.												
-		Ser	Phe	Ala	Gly 20	Val	Ala	Ala	Val	Ser 25	Thr	Thr	Ala	Ser	Ala 30	Lys	Ser
5		Tyr	Ala	Thr 35	Ala	Gly	Ala	Tyr	Ser 40	Thr	Leu	Lys	Thr	Asp 45	Ala	Ala	Thr
		Arg	Asn 50	Val	Ģlu	Ala	Thr	Gly 55	Thr	Asn	Ala	Leu	Tyr 60	Thr	Lys	Pro	Gly
10		Thr 65	Val	Lys	Gly	Ala	Lys 70	Val	Val	Ala	Ser	Lys 75	Ala	Thr	Met	Ala	Lys
15	٠	Leu	Ala	Ser	Ser	Lys 85	Lys	Ser	Ala	Asp	Tyr 90	Phe	Arg	Ala	Tyr	Gly 95	Val
		Lys	Thr	Thr	Asn 100	Arg	Gly	Ser	Val	Tyr 105	Tyr	Arg	Val	Val	Thr 110	Met	Asp
20		Gly	Lys	Tyr 115	Arg	Gly	Tyr	[Val	Tyr 120	Gly	Gly	Lys	Ser	Asp 125	Thr	Ala	Phe
		Ala	Gly 130	Gly	Ile	Lys	Ser	Ala 135	Glu	Thr	Thr	Thr	Lys 140	Ala	Asp	Met	Pro
25		Ala 145	Arg	Thr	Thr	Glÿ	Phe 150	Tyr	Leu	Thr	Asp	Thr 155	Ser	Lys	Asn	Thr	Leu 160
30	:	Trp	Thr	Ala	Pro	Lys 165	Tyr	Thr	Gln	Tyr	Lys 170	Ala	Ser	Lys	Val	Ser 175	Leu
		Tyr	Gly	Val	Ala 180	Lys	Asp	Thr	Lys	Phe 185	Thr	Val	Asp	Gln	Ala 190		Thr
35		Lys	Thr	Arg 195		Gly	Ser	Leu	Tyr 200	Tyr	His	Val	Thr	Ala 205	Thr	Asn	Gly
		Ser	Gly 210	Ile	Ser	Gly	Trp	11e 215	Tyr	Àla	Gly	Lys	Gly 220	Phe	Ser	Thr	Thr
40		Ala 225	Thr	Gly	Thr	Gln	Val 230		Gly	Gly	Leu	Ser 235	Thr	Asp	Lys	Ser	Val 240
45		Thr	Ala	Thr	Asn	Asp 245	Asn	Ser	Val	Lys	Ile 250		Tyr	Arg	Thr	Thr 255	Asp
1,5		Gly	Thr	Gln	Val 260	Gly	Ser	Asn	Thr	Trp 265	Val <sub>.</sub>	Thr	Ser	Thr	Asp 270	Gly	Thr
50		Lys	Ala	Gly 275	Ser	Lys	Val		_	Lys						Ala	
		Glu	Ala 290	Tyr	Ile	Asn	Ala	Asn 295	Lys	Pro	Ser	Gly	Tyr 300	Thr	Val	Thr	Asn
55	ņi -	Pro 305	Asn	Ala	Ala	Asp	Ala 310	Thr	Tyr	Gly	Asn	Thr 315	Val	Tyr	Ala	Thr	Val 320

	\$	Ser	Gln	Ala	Ala	Thr 325	Ser	Lys	Val	Ala	Leu 330	Lys	Val	Ser	Gly	Thr 335	Pro
5	7	/al	Thr	Thr	Ala 340	Leu	Thr	Thr	Ala	Asp 345	Ala	Asn	Asp	Lys	Val 350	Ala	Ala
	; ]	Asn	Asp	Thr 355	Thr	Ala	Asn	Gly	Ser 360	Ser	Val	Ala	Gly	Ser 365	Thr	Val	Tyr
10	1	Ala	Ala 370	Gly	Thr	Lys	Leu	Ala 375	Gln	Leu	Thr	Thr	Asp 380	Leu	Thr	Gly	Glu
15		ys 385	Gly	Gln	Val	Val	Thr 390	Leu	Thr	Ala	Ile	Asp 395	Thr	Asp	Leu	Gľu	Asp 400
1,5		Ala	Thr	Phe	Thr	Gly 405	Thr	Thr	Thr	Tyr	Tyr 410	Ser	Asp	Leu	Gly	Lys 415	Ala
20	. 1	ſyr	His	Tyr	Thr 420	Tyr	Thr	Tyr	Asn	Lys 425	Asp	Ser	Ala	Ala	Ser 430	Ser	Asn
		Ala	Ser	Thr 435	Gln	Phe	Gly	Ser	Asn 440	Val	Thr	Gly	Thr	Leu 445	Thr	Ala	Thr
25		Leu	Val 450	Met	Ġly	Lys	Ser	Thr 455	Ala	Thr	Ala	Asn	Gly 460	Thr	Thr	Trp	Phe
30	. 4	Asn 165	. M N TP 1	CONT E	or s	· · · · · · · ·	ים אכ	1 1 1 1	١.	•							
	•		SEQU	JENCE	CHA	RACI	ERIS	STICS	S :								·
35			(C)	STR	PE: a RANDE POLOG	DNES	S:	id				.:					
40	(i (xi)		•		TYF	-	•		n Mo	. 17							
					Ala	Thr	. •		-	Ser	Val	Lys'	Ile	Val	Tyr		Thr
45		hr	Asp	Gly	Thr	5 Gln	Val	Gly	Ser	Asn	10 Thr	Trp	Val	Thr		15 Thr	Asp
	G	Sly	Thr		20 Ala	Gly	Ser								30 Asp	Gln	Thr
50	<i>P</i> .	la	Leu 50	35 Glu	Ala	Tyr	Ile			 Asn					Tyr	Thr	Val
55		hr 5		Pro	Asn	Ala :	Ala 70		Ala	Thr		Gly 75	Asn	Thr	Val	Tyr	Àla 80 .

	Thr	Val	Ser	Gln	Ala 85	Ala	Thr	Ser	Lys	Val 90	Ala	Leu	Lys	Val	Ser 95	Gly	
5	Thr	Pro	Val	Thr 100	Thr	Ala	Leu	Thr	Thr 105	Ala	Asp	Ala	Asn	Asp 110	Lys	Val	
	Ala	Ala	Asn 115	Asp	Thr	Thr	Ala	Asn 120	Gly	Ser	Ser	Vaļ	Ala 125	Gly	Ser	Thr	
10	Val	Tyr 130	Ala	Ala	Gly	Thṛ	Lys 135	Leu	Ala	Gln	Leu	Thr 140	Thr	Asp	Leu	Thr	
15	Gly 145	Glu	Lys	Gly	Gln	Val 150	Val	Thr	Leu	Thr	Ala 155	Ile	Asp	Thr	Asp	Leu 160	
	Glu	Asp	Ala	Thr	Phe 165		Gly	Thr	Thr	Thr 170	Tyr	Tyr	Ser	Asp	Leu 175	Gly	
20	Lys	Ala	Tyr	His 180	Tyr	Thr	Tyr	Thr	Tyr 185	Asn	Lys	Asp	Ser	Ala 190	Ala	Ser	
	Ser	Asn	Ala 195	Ser	Thr	Gln	Phe	Gly 200	Ser	Asn	Val	Thr	Gly 205	Thr	Leu	Thr	
25	Ala							•		· :		· · · · · · · · · · · · · · · · · · ·				•	÷
	(2) INFO	RMATI	ON F	FOR S	EQ I	D NC	): 14	:							•	·	٠
3 0	(主).		LEN	IGTH :	627	bas	e pa		*				•	·			
· · ·	·	, (C)	TYP STR TOP	ANDE	DNES	s: E											
3.5	(ii)	MOLĖ	CULE	TYP	E: D	NA (	genc	mic)						. "			
	(xi) SEQ	QUENC	E DE	SCRI	PTIC	N: S	EQ I	D NO	): 14	· :	-						
10	TCAGTTACA	AG CA	ACCA	ACGA	. TAA	CAGT	GTT	AAGA	TTGT	TT A	CCGI	'ACGA	C TC	ATGG	CACI	•	60
	CAAGTTGGT	T CT	'AACA	.CTTG	GGT	'AACT	TCA	ACTG	ATGG	TA C	:AAAG	GCAG	G TI	CTAA	.GGTA		120
	AGCGATAAG	G CC	GCCG	ATCA	AAC	TGCT	CTT	GAAG	CCTA	CA T	CAAT	GCTA	A CA	AGCC	TAGO		180
15	GGTTACACT	G TA	ACTA	ACCC	TAA	TGCT	GCA	GATG	CTAC	CT A	TGGT	AACA	.C AG	TTTA	.CGCT		240
	ACTGTTTCC	C AA	GCAG	CTAC	TTC	TAAG	GTC	GCTT	TGAA	GG T	CTCA	GGGA	C TC	CTGT	TACT		300
0	ACTGCATTG	A CT	ACAG	CTGA	TGC	TAAT	GAT	AAGG	TTGC	AG C	TAAC	GATA	C CA	.CTGC	TAAT		360
	GGTAGTTCT	G TT	GCAG	GCTC	AAC	AGTC	TAT	GCTG	CTGG	TA C	TAAG	TTGG	C TC	AATT	AACA		420
5	ACTGACTTG	A CT	GGTG	AAAA	GGG	TCAA	GTT	GTCA	CATT.	AA C	TGCC	ATCG	A TA	CTGA	TTTG		480

GAAGACGCTA CGTTCACTGG AACTACGACT TACTATTCAG ATCTTGGTAA AGCATACCAC

	TACACTTACA CTTACAATAA GGACAGTGCT GCTTCTTCAA ATGCAAGTAC CCAATTTGGT	600
	TCAAACGTCA CTGGTACTTT AACTGCT	627
5	(2) INFORMATION FOR SEQ ID NO: 15:	
10	<ul> <li>(i) SEQUENCE CHARACTERISTICS:         <ul> <li>(A) LENGTH: 1680 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: both</li> </ul> </li> </ul>	
	(D) TOPOLOGY: both	
	(ii) MOLECULE TYPE: DNA (genomic)	
15	(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION:1891583	
20	(ix) FEATURE: (A) NAME/KEY: misc_signal (B) LOCATION:189278	. *
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:	
	TCCAACGACA ATCAGAGCGT AATCCTTGTA TCTCCTTAAG GAAATCGCTA TACTTATCTT	60
	CGTAGTTAGG GGATAGCTGA TCGGGTCCGC TAATGTTATG AAATAAAATT CTTAACAAAA	120
30	GCGCTAACTT CGGTTATACT ATTCTTGCTT GATAAATTAC ATATTTTATG TTTGGAGGAA	180
	GAAAGATT ATG CAA TCA AGT TTA AAG AAA TCT CTT TAC TTG GGC CTT GCC  Met Gln Ser Ser Leu Lys Lys Ser Leu Tyr Leu Gly Leu Ala  1 5 10	230
35	GCA TTG AGC TTT GCT GGT GTT GCT GCC GTT TCA ACG ACT GCT TCA GCT	278
	Ala Leu Ser Phe Ala Gly Val Ala Ala Val Ser Thr Thr Ala Ser Ala  15 20 25 30	
40	AAG TCA TAC GCT ACT GCA GGT GCC TAT TCA ACG TTA AAG ACG GAC GCT Lys Ser Tyr Ala Thr Ala Gly Ala Tyr Ser Thr Leu Lys Thr Asp Ala 35 40 45	326
45	GCT ACT CGT AAC GTC GAA GCT ACT GGT ACT AAC GCT TTA TAC ACG AAG Ala Thr Arg Asn Val Glu Ala Thr Gly Thr Asn Ala Leu Tyr Thr Lys 50 55 60	374
	CCA GGT ACT GTT AAG GGT GCT AAG GTT GTC GCT TCT AAG GCT ACT ATG	422
5 Q	Pro Gly Thr Val Lys Gly Ala Lys Val Val Ala Ser Lys Ala Thr Met 65 70 75 75	
55	GCT AAG TTA GCT TCT TCA AAG AAG TCA GCT GAC TAC TTC CGT GCT TAC Ala Lys Leu Ala Ser Ser Lys Lys Ser Ala Asp Tyr Phe Arg Ala Tyr 80 85 90	470
٠ بـ	GGT GTT AAG ACC ACT AAC CGT GGT TCA GTT TAC TAC CGT GTT GTA ACG Gly Val Lys Thr Thr Asn Arg Gly Ser Val Tyr Tyr Arg Val Val Thr	518

	99	5				100	)				10	5 .				110		
5						Arg					r Gl					C ACT Thr		566
10	Ala	TT a Pho	T GCT e Ala	r GG a Gly 130	y Gly	T ATO	AAG Lys	TCT Ser	GCT Ala 135	ı Glu	A ACC	G ACT	r ACT	AAC Lys	Ala	GAT Asp		614
	ATC			a Arc					Туг					Ser		AAC Asn		662
15			ı Trp					Tyr					Ala			GTT Val		710
20		Let					Lys					Thr	GTA Val			GCT Ala 190		758
25	GCT Ala	ACT Thr	AAG Lys	ACI Thr	CGT Arg 195	GAA Glu	GGT Gly	TCA Ser	Leu	TAC Tyr 200	Tyr	CAC His	GTA Val	ACT Thr	GCT Ala 205	ACT Thr		806
30					Ile								AAG Lys					854
													TCA Ser 235					902
35	TCA Ser	GTT Val 240	Thr	GCA Ala	ACC Thr	AAC Asn	GAT Asp 245	AAC Asn	AGT Ser	GTT Val	AAG Lys	ATT Ile 250	GTT Val	TAC Tyr	CGT Arg	ACG Thr		950
40	ACT Thr 255	GAT Asp	GGC Gly	ACT Thr	CAA Gln	GTT Val 260	GGT Gly	TCT Ser	AAC Asn	ACT Thr	TGG Trp 265	GTA Val	ACT	TCA Ser	ACT Thr	GAT Asp 270	·	998
45	GGT Gly	ACA Thr	AAG Lys	GCA Ala	GGT Gly 275	TCT Ser	AAG Lys	GTA Val	AGC Ser	GAT Asp 280	AAG Lys	GCC Ala	GCC Ala	GAT Asp	CAA Gln 285	ACT Thr		1046
50	GCT Ala	CTT	GAA Glu	GCC Ala 290	TAC Tyr	ATC Ile	AAT Asn	GCT Ala	AAC Asn 295	AAG Lys	CCT Pro	AGC Ser	GGT Gly	TAC Tyr 300	ACT Thr	GTA Val		1094
	ACT Thr	AAC Asn	CCT Pro 305	AAT Asn	GCT Ala	GCA Ala	Asp	GCT Ala 310	ACC Thr	TAT Tyr	GGT Gly	AAC Asn	ACA Thr 315	GTT Val	TAC Tyr	GCT Ala		1142
55	Thr	GTT Val 320	Ser	CAA Gln	GCA Ala	GCT Ala	ACT Thr 325	TCT Ser	AAG Lys	GTC Val	GCT Ala	TTG Leu 330	AAG Lys	GTC Val	TCA Ser	GGG Gly		1190
			5															

	ACT	CCT	GTT	ACT	ACT	GCA	TTG	ACT	ACA	GCT	GAT	GCT	AAT	GAT	AAG	GTT	•	1238
	Thr 335	Pro <sub>.</sub>	Val	Thr	Thr	Ala 340	Leu	Thr	Thr	Ala	Asp 345	Ala	Asn	Asp	Lys	350		
5 .															TCA Ser 365	Thr		1286
10															TTG Leu			1334
15															GAT			1382
20															CTT Leu			143
20																TCT Ser 430		1478
25															TTA Leu 445			1526
30															ACT Thr			1574
3,5		TTC Phe		TAA	PAAT'	TAT :	ratt:	ragg:	rg Ao	GCTT	rgtT(	G AT	<b>LAAA</b> A	AGGT			:	1623
	CTT	rtcai	ACG T	rtta:	rgtt	GG GC	BAGA(	CCTT	r TT2	ATAT	rgaa	AAA	ATTA	GGC (	CTTT(	3TT		1680
40	(2)	INF	CAMAC	rion	FOR	SEQ	ID 1	10: 3	16:					·				
45		·	(2 (E	A) LI B) T	engti YPE :	1: 46 amii	RACTI 55 ar no ac line	mino cid			٠							
•		(ii	) MOI	LECUI	LE T	YPE:	prot	ein										
50				_			[PTIO								<u>.</u>			
	Met 1	Gln	Ser	Ser	Leu 5	Lys	Lys	Ser		Tyr 10	Leu	Gly	Leu	Ala	Ala 15	Leu		
55	Ser	Phe	Ala	Gly 20	Val	Ala	Ala	Val	Ser 25	Thr	Thr	Ala	Ser	Ala 30	Lys	Ser		
	Т12~	ר ו א	Th ~	77.	Clv	<b>71</b> -	T1	Car	ጥከ፦	T.An	Lvc	Thr	y e.c.	Δla	Δla	Thr		

40 35 Arg Asn Val Glu Ala Thr Gly Thr Asn Ala Leu Tyr Thr Lys Pro Gly 5 Thr Val Lys Gly Ala Lys Val Val Ala Ser Lys Ala Thr Met Ala Lys Leu Ala Ser Ser Lys Lys Ser Ala Asp Tyr Phe Arg Ala Tyr Gly Val 10 Lys Thr Thr Asn Arg Gly Ser Val Tyr Tyr Arg Val Val Thr Met Asp 15 Gly Lys Tyr Arg Gly Tyr Val Tyr Gly Gly Lys Ser Asp Thr Ala Phe 115 Ala Gly Gly Ile Lys Ser Ala Glu Thr Thr Thr Lys Ala Asp Met Pro 135 140 20 Ala Arg Thr Thr Gly Phe Tyr Leu Thr Asp Thr Ser Lys Asn Thr Leu 150 145. Trp Thr Ala Pro Lys Tyr Thr Gln Tyr Lys Ala Ser Lys Val Ser Leu 170 25 Tyr Gly Val Ala Lys Asp Thr Lys Phe Thr Val Asp Gln Ala Ala Thr 180 Lys Thr Arg Glu Gly Ser Leu Tyr Tyr His Val Thr Ala Thr Asn Gly Ser Gly Ile Ser Gly Trp Ile Tyr Ala Gly Lys Gly Phe Ser Thr Thr 215 35 Ala Thr Gly Thr Gln Val Leu Gly Gly Leu Ser Thr Asp Lys Ser Val 230 Thr Ala Thr Asn Asp Asn Ser Val Lys Ile Val Tyr Arg Thr Thr Asp 40 Gly Thr Gln Val Gly Ser Asn Thr Trp Val Thr Ser Thr Asp Gly Thr 265 260 Lys Ala Gly Ser Lys Val Ser Asp Lys Ala Ala Asp Gln Thr Ala Leu 45 280 Glu Ala Tyr Ile Asn Ala Asn Lys Pro Ser Gly Tyr Thr Val Thr Asn 290 295 50 Pro Asn Ala Ala Asp Ala Thr Tyr Gly Asn Thr Val Tyr Ala Thr Val 310 Ser Gln Ala Ala Thr Ser Lys Val Ala Leu Lys Val Ser Gly Thr Pro 55 330

Val Thr Thr Ala Leu Thr Thr Ala Asp Ala Asn Asp Lys Val Ala Ala

				340					345					350		
5	Asn	Asp	Thr 355	Thr	Ala	Asn	Gly	Ser 360	Ser	Val	Ala	Gly	Ser 365	Thr	Val	Tyr
	Ala	Ala 370	Gly ·	Thr	Lys	Leu	Ala 375	Gln	Leu	Thr	Thr	Asp 380	Leu	Thr	Gly	Glu
10	Lys 385	Gly	Gln	Val	Val	Thr 390	Leu	Thr	Ala	Ile	Asp 395	Thr	Asp	Leu	Glu	Asp 400
·	Ala	Thr	Phe	Thr	Gly 405	Thr	Thr	Thr	Tyr	Tyr 410	Ser	Asp	Leu	Gly	Lys 415	Ala
15	Tyr	His	Tyr	Thr 420	Tyr	Thr	Tyr	Aşn	Lys 425	Asp	Ser	Ala	Ala	Ser 430	Ser	Asn
20	Ala	Ser	Thr 435	Gln	Phe	Gly	Ser	Asn 440	Val	Thr	Gly	Thr	Leu 445	Thr	Ala	Thr
	Leu	Val 450	Met	Gly	Lys	Ser	Thr 455	Ala	Thr	Ala	Asn	Gly 460	Thr	Thr	Trp	Phe
25	Asn 465	. ·						-	٠.					.*		

#### What we claim is:

- 1. A DNA molecule encoding a polypeptide capable of binding to human and/or animal epithelial cell types, said DNA molecule being selected from the group consisting of:
  - (a) DNA molecules having at least the partial coding sequences of any one of SEQ ID NO. 1, SEQ ID NO. 2, SEQ ID NO. 3, SEQ ID NO. 4, SEQ ID NO 5 and/or SEQ ID NO. 6 excluding the full length SEQ ID NO 6;
- (b) DNA molecules encoding a polypeptide having at least the partial amino acid sequences of any one of SEQ ID NO. 7, SEQ ID NO. 8, SEQ ID NO.9, SEQ ID NO. 10, SEQ ID NO. 11 and/or SEQ ID NO 12, excluding the full length SEQ ID NO 12;
- (c) DNA molecules the coding sequences of which differ from the coding sequence of a nucleic acid molecule of (a) or (b) due to the degeneracy of the genetic code;
- (d) DNA molecules hybridizing under stringent conditions to a molecule of (a), (b) and/or (c) excluding the full-length SEQ ID NO 6; and
  - (e) DNA molecules encoding a polypeptide capable of binding to human and/or animal epithelial cell types and having an amino acid sequence which shows at least 40 % identity to a sequence contained in (b) excluding the full length SEQ ID NO 12.
- 2. The DNA molecule of claim 1 encoding a polypeptide capab
  - 2. The DNA molecule of claim 1 encoding a polypeptide capable of binding to human and/ or animal epithelial cell types, wherein said polypeptide is capable of binding to intestinal, urogenital and/or endothelial cells.
- 3. The DNA molecule of claim 1 or 2, which originates from the DNA molecule encoding Lactobacillus brevis S-layer SlpA protein.
  - 4.A vector containing a DNA molecule of any one of claim 1, 2 or 3.
- 5. The vector of claim 4, in which the DNA molecule is operably linked to expression and optionally to secretion control sequences allowing expression in prokaryotic or eukaryotic

host cells.

6.A host cell transformed with a DNA molecule of any one of claims 1, 2 or 3 or with a vector of claim 4 or 5.

- 7. A method of constructing a host cell capable of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cell types, comprising transforming the cell with at least one DNA molecule selected from the group consisting of:
- 10 (a) DNA molecules having at least the partial coding sequences of any one of SEQ ID NC 1, SEQ ID NO. 2, SEQ ID NO. 3, SEQ ID NO. 4, SEQ ID NO 5, and/or SEQ ID NO 6;
  - (b) DNA molecules encoding a polypeptide having at least the partial amino acid sequences of any one of SEQ ID NO. 7, SEQ ID NO. 8, SEQ ID NO. 9, SEQ ID NO.10, SEQ ID NO. 11 and/or SEQ ID NO 12;
- 15 (c) DNA molecules, the coding sequence of which differ from the coding sequence of a DNA molecule of (a) or (b) due to the degeneracy of the genetic code;
  - (d) DNA molecules hybridizing under stringent conditions to a molecule of (a), (b) and/or (c); and
- (e) DNA molecules encoding a polypeptide capable of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cell types and having an amino acid sequence which shows at least 40 % identity to a sequence contained in (b).
  - 8. A host cell constructed by the method of claim 7.
- 9. The host cell of any one of claims 6 or 8, which has probiotic effects.
  - 10. The host cell of claim 9, wherein the probiotic effects have been enhanced by genetic means.
- 11. The host cell of claim 9 or 10, which belongs to lactic acid bacteria or bifidobacteria.
  - 12. The host cell of any one of claims 6 or 8 to 11, wherein the host cell is a vaccine carrier.

13. The host cell of any one of claims 6 or 8 to 12, wherein the host cell has been genetically modified to carry at least one of the factors selected from the group consisting of an antigen, an epitope, a single chain antibody, an enzyme, a pharmaceutical protein and a toxin.

5

- 14. A method of constructing a polypeptide capable of binding to human and/or animal epithelial cell types, like intestinal, urogenital and/or endothelial cell types, comprising modifying the gene encoding the polypeptide with a DNA molecule of claim 1 or 7.
- 10 15. A gene encoding a preselected protein, wherein the gene encoding the protein is genetically modified to bind to human and/or animal epithelial cell types, like intestinal, urogenital and/ or endothelial cell types with at least one DNA molecule selected from the group consisting of any of the sequences of claim 1 or 7.
- 16. The gene of claim 15, wherein the preselected protein is selected from the group consisting of an antigen, an epitope, a single chain antibody, an enzyme, a pharmaceutical protein and a toxin.
  - 17. A polypeptide encoded by a DNA molecule of claim 1 or 15.

- 18. A Lactobacillus brevis strain for carrying preselected factors/ properties to human and/or animal epithelial cells or cell surfaces, wherein the strain has the capability of binding to human and/or animal epithelial cell types and wherein the gene encoding the S-layer SlpA protein is genetically modified to carry at least one of the factors/properties selected from the group consisting of an antigen, an epitope, a single chain antibody, an enzyme, a pharmaceutical protein and a toxin.
- 19. A Lactobacillus brevis strain for carrying preselected factors/ properties to human and/or animal epithelial cells or cell surfaces, wherein the strain has the capability of binding to human and/or animal epithelial cell types and wherein the strain is genetically modified to have improved probiotic effects.

- 20. A Lactobacillus brevis S-layer SlpA protein for carrying preselected factors/ properties to human and/or animal epithelial cells or cell surfaces, wherein the protein has the capability of binding to human and/or animal epithelial cell types and wherein the gene encoding the S-layer SlpA protein is genetically modified to carry at least one of the factors/properties selected from the group consisting of an antigen, an epitope, a single chain antibody, an enzyme, a pharmaceutical protein and a toxin.
- 21. A method for carrying preselected factors/ properties to human and/or animal epithelial cells or cell surfaces by using a *Lactobacillus brevis* strain, wherein the strain has the capability of binding to human and/or animal epithelial cell types and wherein the gene encoding the S-layer SlpA protein is genetically modified to carry at least one of the factors/properties selected from the group consisting of an antigen, an epitope, a single chain antibody, an enzyme, a pharmaceutical protein and a toxin.
- 22. A method for carrying preselected factors/ properties to human and/or animal epithelial cells or cell surfaces by using a *Lactobacillus brevis* strain, wherein the strain has the capability of binding to human and/or animal epithelial cell types and wherein the strain is genetically modified to have improved probiotic effects.
- 23. A method for carrying preselected factors/ properties to human and/or animal epithelial cells or cell surfaces by using the *Lactobacillus brevis* S-layer SlpA protein, wherein the protein has the capability of binding to human and/or animal epithelial cell types and wherein the gene encoding the S-layer SlpA protein is genetically modified to carry at least one of the factors/properties selected from the group consisting of an antigen, an epitope, a single chain antibody, an enzyme, a pharmaceutical protein and a toxin.
  - 24. The use of the host cells of any one of claims 6 or 8 to 13 for excluding pathogens in the cell surfaces of the gastrointestinal or urogenital tract of humans and/ or animals.

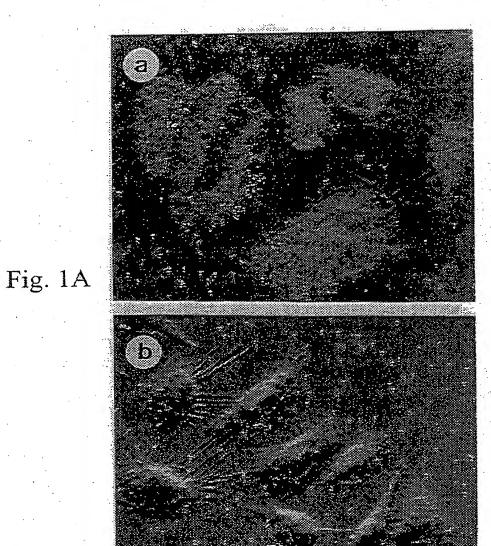


Fig. 1B

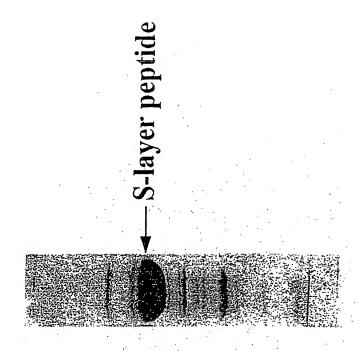
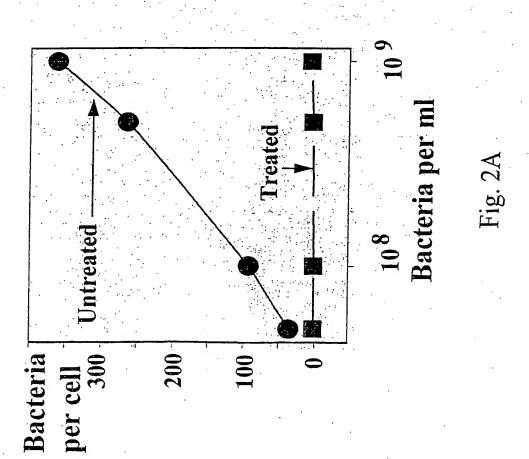
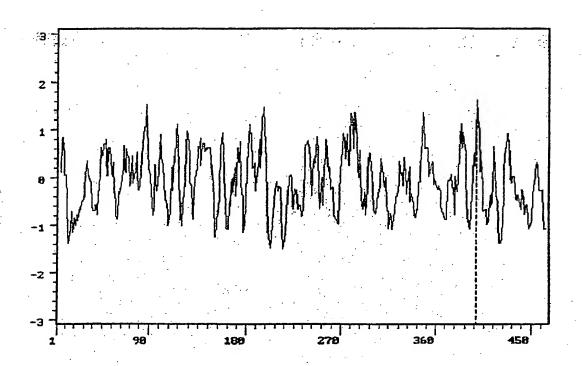


Fig. 2B



SUBSTITUTE SHEET (Rule 26)

# Expression of slpA fragments as fusions to fliC



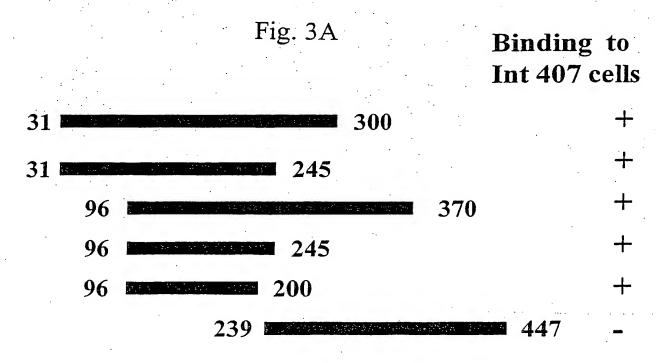
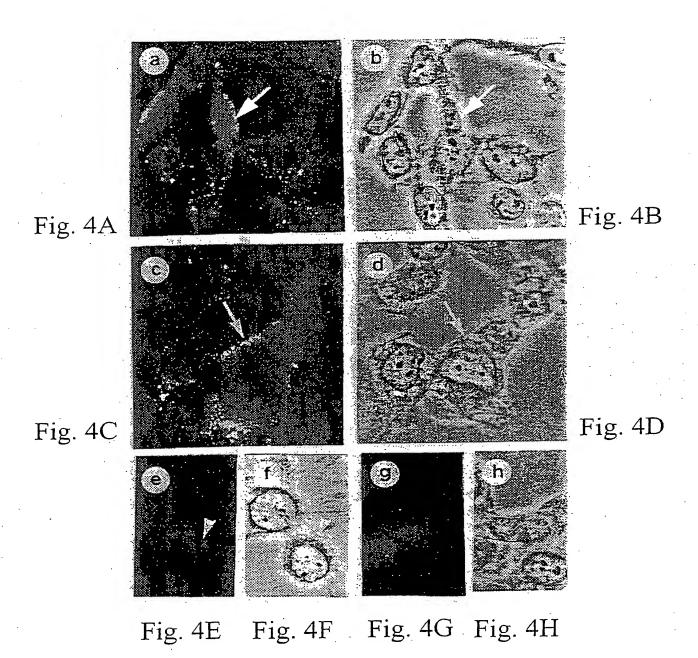


Fig. 3B



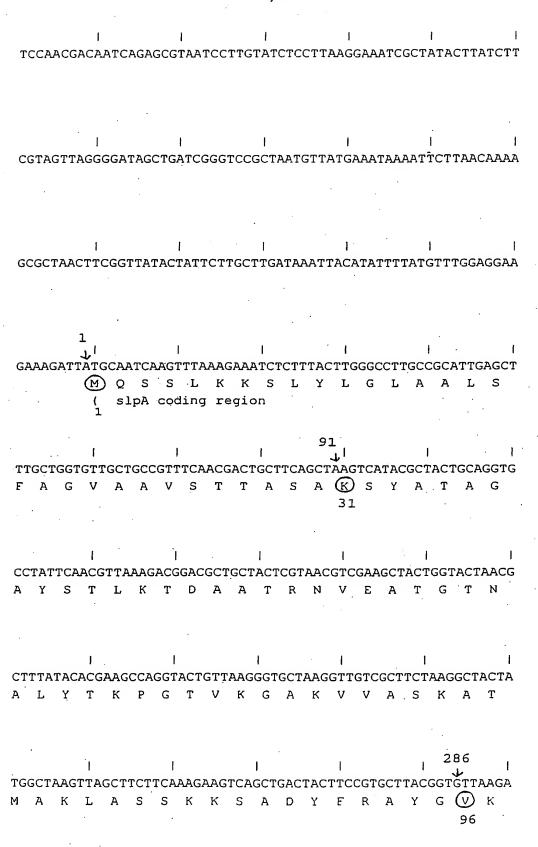


Fig. 5A

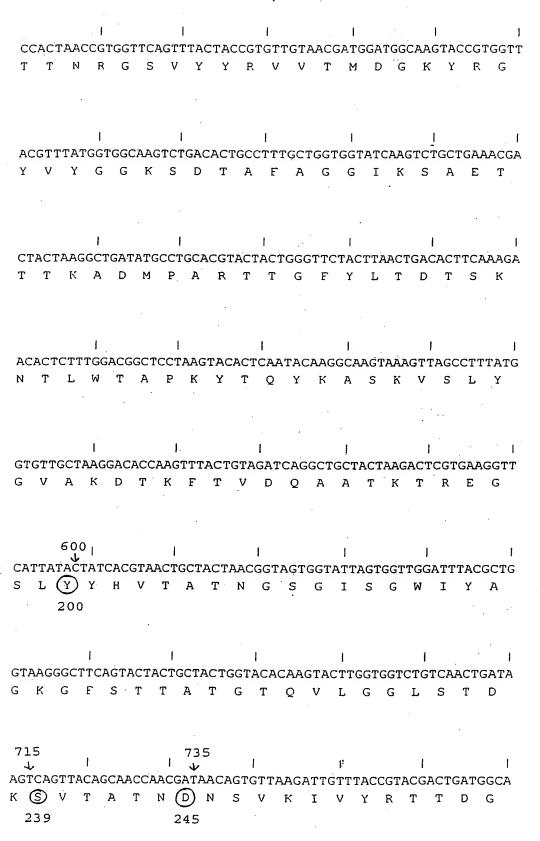


Fig. 5B

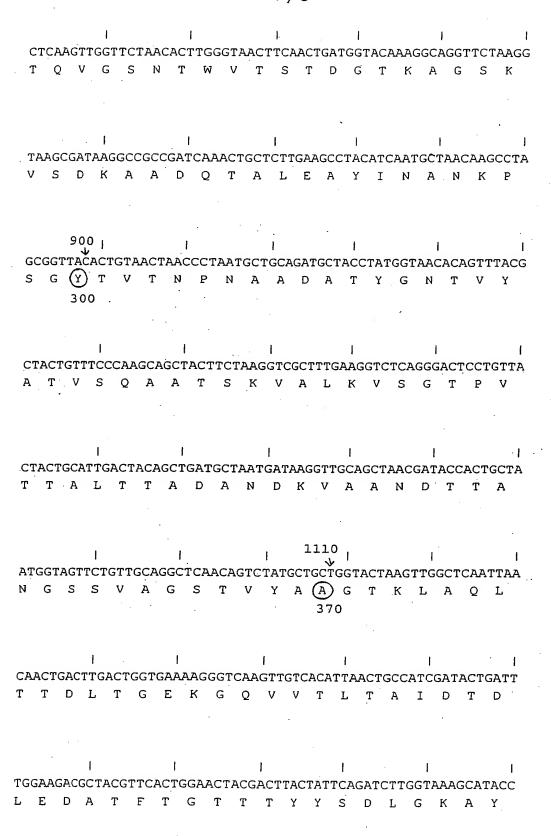


Fig. 5C

8/8

Fig. 5D

#### INTERNATIONAL SEARCH REPORT

International application No. PCT/FI 99/00290

#### A. CLASSIFICATION OF SUBJECT MATTER IPC6: C07K 14/325 According to International Patent Classification (IPC) or to both national classification and IPC B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC6: CO7K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched SE,DK,FI,NO classes as above Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) WPI, PATENT ABSTRACTS OF JAPAN, STRAND, CASEARCH, BIOSIS, MEDLINE, SCISEARC H, LIFE SCIENCES COLLECTION C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. FEMS Microbiology Reviews, Volume 20, 1997, Hubert Bahl et al, "IV. Molecular biology of Х 18-23 S-layers", page 47 - page 98, see pages 82-83 A 1-17,24WO 9400581 A1 (VIAGEN OY), 6 January 1994 1 - 24A (06.01.94)Journal of Bacteriology, Volume 174, No 22, 1992, 1 - 24Gabriele vidgrén et al, "S-Layer Protein Gene of Lactobacillus brevis: Cloning by Polymerase ChainReaction and Determinaiton of the Nucleotide Sequence" page 7419 - page 7427 X Further documents are listed in the continuation of Box C. See patent family annex. later document published after the international filing date or priority Special categories of cited documents: date and not in conflict with the application but cited to understand "A" document defining the general state of the art which is not considered the principle or theory underlying the invention to be of particular relevance "E" erher document but published on or after the international filing date document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other step when the document is taken alone special reason (as specified) document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination document referring to an oral disclosure, use, exhibition or other being obvious to a person skilled in the art document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 2 8 -07- 1999 26 July 1999 Name and mailing address of the ISA/ Authorized officer Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Hampus Rystedt/EÖ Facsimile No. +46 8 666 02 86 +46 8 782 25 00 Telephone No.

### INTERNATIONAL SEARCH REPORT

International application No. PCT/FI 99/00290

	PC1/	/FI 99/0	0290
C (Continu	ation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant pa	issages	Relevant to claim N
A	Journal of Applied Bacteriology, Volume 74, 1993, C. Schneitz et al, "Adhesion of Lactobacillusacidophilus to avian intestinal epithel ial cells mediated by the crystalline bacterial cell surface layer (S-layer)" page 290 - page 294	-	1-24
	<del></del>		
A .	Chemical Abstracts, Volume 126, No 13, 31 March 1997 (31.03.97), (Columbus, Ohio, USA Savijoki, Kirsi et al, "High level heterologou protein production in Lactococcus and Lactobacillus using a new secretion system bas on the Lactobacillus brevis S-layer signals", 1, THE ABSTRACT No 167032, Gene 1997, 186 (2),	ed page	1-24
	255-262		
•			
			•
			•
	+		
	4 .:		
			•
			·
	<u>a andre de la companya de la compan</u>	. A	and the second

### INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

01/07/99

PCT/FI 99/00290

Patent document cited in search report Publication date Patent family member(s) Publication date

WO 9400581 A1 06/01/94 AU 4329793 A 24/01/94

Form PCT/ISA/210 (patent family annex) (July 1992)

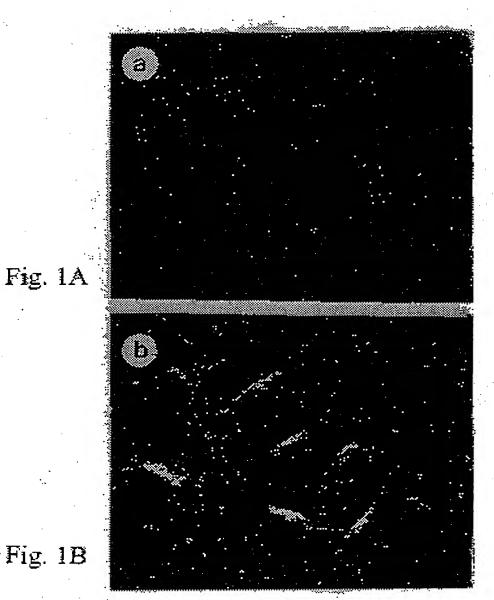


Fig. 1B

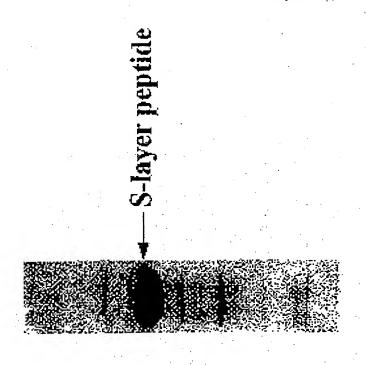
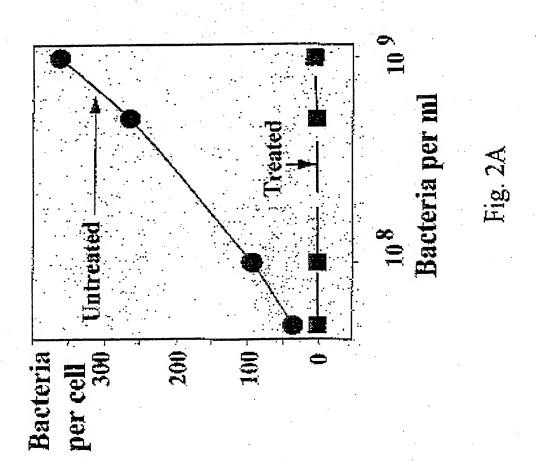
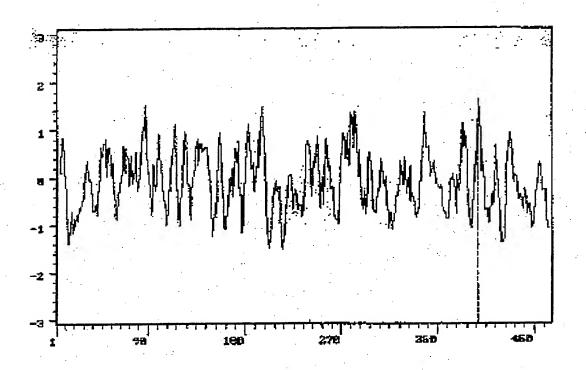


Fig. 2B



SUBSTITUTE SHEET (Rule 26)

# Expression of slpA fragments as fusions to fliC



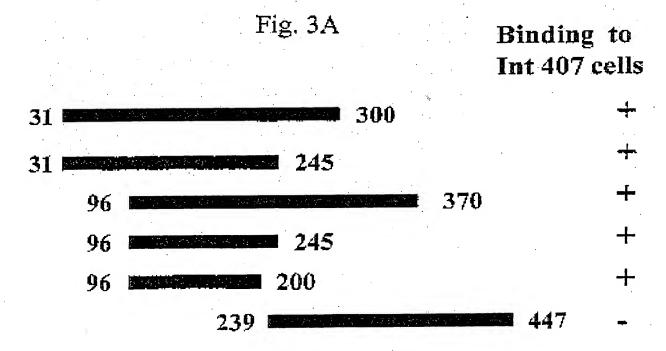
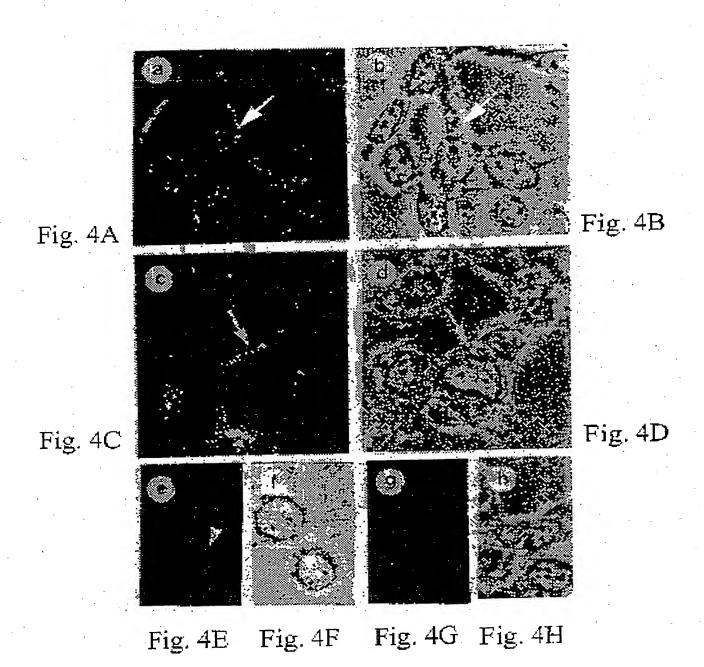


Fig. 3B



SUBSTITUTE SHEET (Rule 26)

5/8

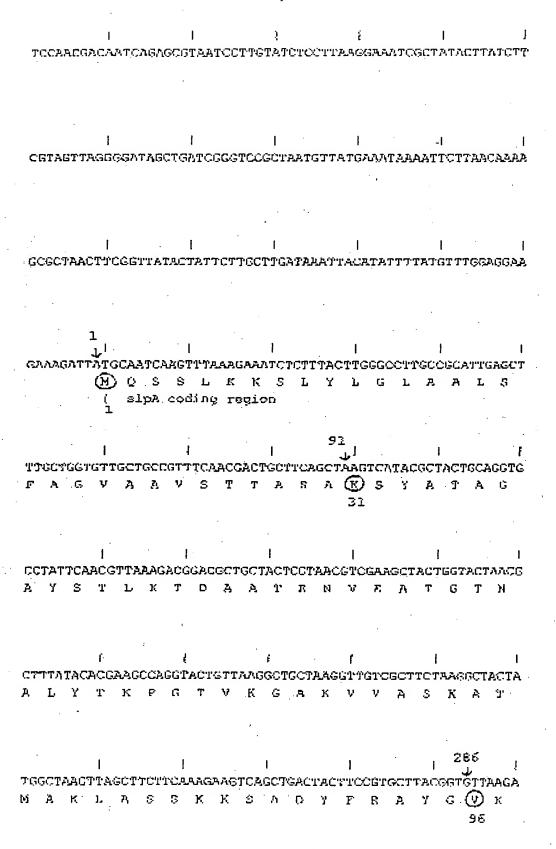


Fig. 5A

t is the best of the partition

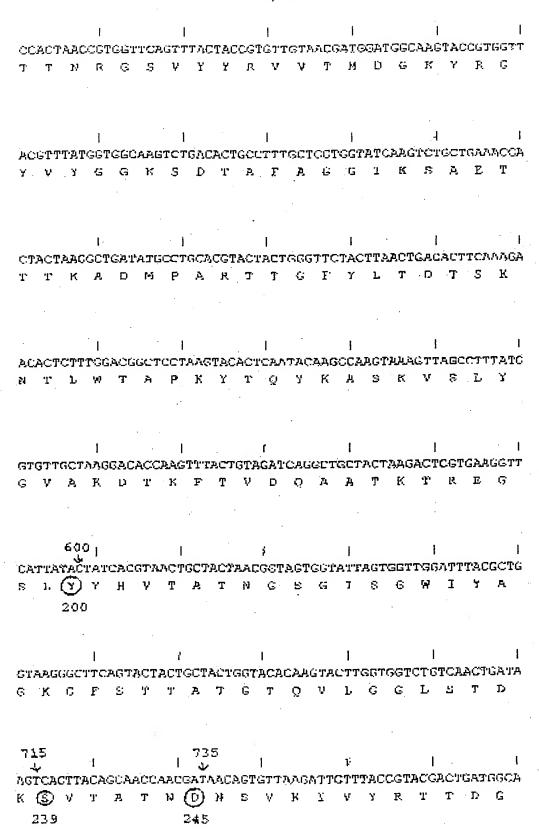


Fig. 5B

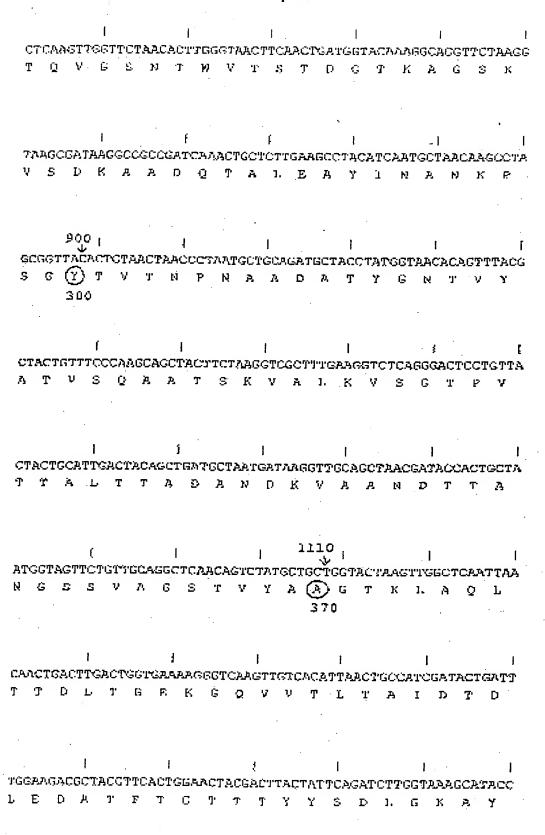


Fig. 5C

ξ.

8/8

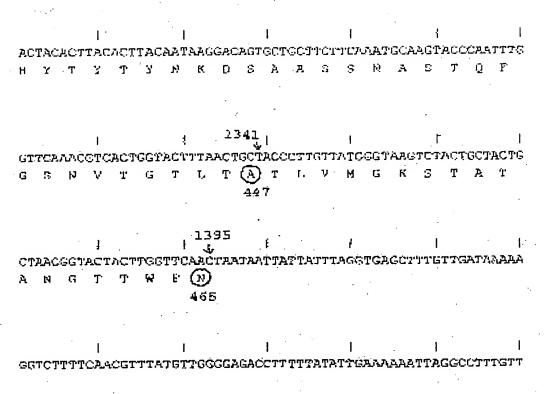


Fig. 5D

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
$\square$ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
OTHER:

## IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.